Europäisches Patentamt European Patent Office

Office européen des brevets



(11) EP 0 733 709 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication: 25.09.1996 Bulletin 1996/39

(21) Application number: 95115423.6

(22) Date of filing: 29.09.1995

(51) Int. Cl.⁶: **C12N 15/54**, C12N 9/10, C12N 15/70, C12N 1/21

(84) Designated Contracting States: BE CH DE FR GB IT LI SE

(30) Priority: 14.02.1995 JP 25253/95

(71) Applicant: TOYOTA JIDOSHA KABUSHIKI KAISHA Aichi-ken (JP)

(72) Inventors:

Ayumi, Koike,
 c/o Toyota Jidosha K.K.
 Toyota-shi, Alchi-ken (JP)

Tokuzo, Nishino,
 c/o Toyota Jidosha K.K.
 Toyota-shi, Aichi-ken (JP)

Shusei, Obata,
 c/o Toyota Jidosha K.K.
 Toyota-shi, Aichi-ken (JP)

 Shinichi, Ohnuma, c/o Toyota Jidosha K.K.
 Toyota-shi, Aichi-ken (JP) Takeshi, Nakazawa, c/o Toyota Jidosha K.K. Toyota-shi, Aichi-ken (JP)

Kyozo, Ogura,
 c/o Toyota Jidosha K.K.
 Toyota-shi, Alchi-ken (JP)

Tanetoshi, Koyama,
 c/o Toyota Jidosha K.K.
 Toyota-shi, Aichi-ken (JP)

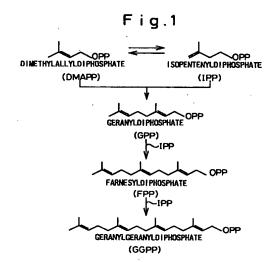
(74) Representative: Tiedtke, Harro, Dipl.-Ing. Patentanwaltsbüro
Tiedtke-Bühling-Kinne & Partner
Bavariaring 4
80336 München (DE)

Remarks:

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Mutated farmesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate and gene coding therefor

(57) A mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate and gene coding for said mutated enzyme, wherein the mutated enzyme is modified from a native farnesyldiphosphate synthase by mutation of a gene coding for a native farnesyldiphosphate synthase.



Description

5

10

45

BACKGROUND OF INVENTION

1. Field of Invention

The present invention relates to the mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate and a process for production thereof, as well as genes coding for said mutated enzymes and a process for isolation thereof.

2. Related Art

In nature there are various isoprenoid chain compounds comprising 5 carbon atom-basic structure, isoprene units, and these isoprenoid compounds play important roles for the life of various organisms. It is known that the chain-extension mechanism is catalyzed by a series of prenyltransferases which catalyze a series of catalytic reactions comprising sequential condensation of isopentenyldiphosphate (IPP) having 5 carbon atoms with its isomer dimethylallyldiphosphate (DMAPP). Among the isoprenoid compounds, farnesyldiphosphate (FPP) having 15 carbon atoms is positioned at a branching point in a biosynthesis pathway, from which various physiologically important start to geranylgeranyldiphosphate (GGPP) having 20 carbon atoms, to quinones, squalene, to steroids, farnesylated protein, dolichol etc.

Different prenyltransferases synthesize different isoprenoid compounds having different lengths. However, prenyltransferases have a common activity to condense an isoprenoid unit to extend the chain, and in fact, amino acids essential for the condensation are being clarified on the basis of homology of amino acid sequences of different prenyltransferases. However, the mechanism which determines the length of the isoprenoid compound have not yet clarified.

A biosynthesis pathway for geranyldiphosphate (GPP), farnesyldiphosphate (FPP) and geranylgeranyldiphosphate (GGPP) starting from an isoprenoid unit is shown in Fig. 1. In this biosynthesis pathway, the prenyltransferase which synthesizes farnesyldiphosphate is designated "farnesyldiphosphate synthase", and the prenyltransferase which synthesizes geranylgeranyldiphosphate is designated "geranylgeranyldiphosphate synthase".

Farnesyldiphosphate synthases are known in <u>Bacillus thermophils</u> (J. Biochem. <u>113</u>, 355 - 363 (1993)), <u>E. coli</u> (J. Biochem. <u>108</u>, 995 - 1000 (1990)), yeast (J.B.C. <u>265</u>, 19176 - 19184 (1989)), rats (Mol. Cell. Biol. <u>7</u>, 3138 - 3146 (1987)) and in humans (J.B.C. <u>265</u>, 4607 - 4616 (1990)), and their amino acid sequences are also known.

On the other hand, geranylgeranyldiphosphate synthases are known in <u>Rhodopseudomonas capusulata</u> (J. Bacteriol. <u>154</u>, 580 - 590 (1983)), <u>Erwinia uredovora</u> (J. Bacteriol. <u>172</u>, 6704 - 6712 (1990)), <u>Sulfolobus acidocaldarius</u> (J.B.C. <u>269</u>, 14792 - 14797 (1994)) etc.

However, it had not been known that an enzyme having geranylgeranyldiphosphate synthase activity can be obtained by mutation of farnesyldiphosphate synthase.

SUMMARY OF INVENTION

Accordingly, the present invention provides a novel geranylgeranyldiphosphate synthase obtainable by mutating a farnesyldiphosphate synthase and a process for production thereof, as well as gene system therefor and a process for isolation of the gene.

More specifically, the present invention provides a process for production of a gene coding for geranylgeranyldiphosphate synthase comprising the steps of:

- (1) subjecting genes coding for a farnesyldiphosphate synthase to a mutagenesis;
- (2) expressing the genes subjected to the mutagenesis, and
- (3) selecting a gene which provides a geranylgeranyldiphosphate synthase.

The present invention further provides a gene coding for geranylgeranyldiphosphate synthase, an expression vector containing said gene, and a host transformed with said vector.

The present invention also provides a process for production of geranylgeranyldiphosphate synthase comprising expressing said gene, and geranylgeranyldiphosphate synthase obtainable by said process.

From another point of view, the present invention provides a geranylgeranyldiphosphate synthase having an amino acid sequence modified from an amino acid sequence of native farnesyldiphosphate synthase wherein the modification is deletion of one or more amino acids, addition of one or more amino acids, and/or replacement of one or more amino acids with other amino acids.

The present invention still further provides a gene coding for the above-mentioned geranylgeranyldiphosphate synthase, a vector, especially an expression vector comprising said gene, and a host transformed with said vector.

The present invention further provides a process for production of geranylgeranyldiphosphate synthase comprising the steps of cultivation said host, and purification the geranylgeranyldiphosphate synthase from the culture.

The present invention further provides a process for production of geranylgeranyldiphosphate or geranylgeranyol, comprising the steps of acting the present geranylgeranyldiphosphate synthase on isopentenyldiphosphate, dimethylallyldiphosphate, geranyldiphosphate or farnesyldiphosphate as a substrate.

BRIEF EXPLANATION OF DRAWINGS

5

10

15

20

25

30

Figure 1 represents a biosynthesis pathway for farnesyldiphosphate and geranylgeranyldiphosphate.

Fig. 2 shows the homology of amino acid sequences of farnesyldiphosphate synthase derived from different species. In this Figure, the sequences in the boxes A to E show regions having relatively high homology and which are expected to participate in enzyme activity.

Fig. 3 shows the homology of amino acid sequences of farnesyldiphosphate synthase derived from different species. In this Figure, the sequences in the boxes F and G show regions having relatively high homology and which are expected to participate in enzyme activity.

Fig. 4 shows a native amino acid sequence of farnesyldiphosphate synthase derived from <u>Bacillus stearother-mophilus</u> (indicated as W.T), and the mutated points in amino acid sequences of the modified enzymes having geranyldiphosphate synthase activity (No. 1 to No. 4).

Fig. 5 schematically shows a process for construction of the present modified gene.

Fig. 6 is a profile of reversed phase TLC (developer: acetone/water = 9/1) showing products formed by acting the present enzyme on a substrate dimethylallyldiphosphate.

Fig. 7 is a profile of a reversed-phase TLC (developer: acetone/water = 9/1) showing products formed by acting the present enzyme on a substrate geranyldiphosphate.

Fig. 8 is a profile of a reversed phase TLC (developer: acetone/water = 9/1) showing products formed by acting the present enzyme on a substrate (all-E)-farnesyldiphosphate.

Fig. 9 is a profile of a reversed phase TLC (developer: acetone/water = 9/1) showing products formed by acting the present enzyme on a substrate (all-E)-farnesyldiphosphate.

DETAILED DESCRIPTION

Genes of the present invention can be obtained by subjecting a gene coding for a farnesyldiphosphate synthase to mutagenesis, expressing the genes subjected to the mutagenesis, and selecting a gene providing a protein having geranylgeranyldiphosphate synthase activity.

Genes coding for a farnesyldiphosphate synthase used in the present invention may be those of any origin. For example, farnesyldiphosphate synthases of <u>E. coli</u>, yeast, human, rat etc., as well as genes coding therefor are known, and amino acid sequences of these enzymes have high homology as shown in Fig. 2. Therefore, in addition to the gene derived from <u>Bacillus stearothermophilus</u> as described in detail, according to the present invention, any gene coding for an amino acid sequence having a high homology, for example, at least 20% homology with the amino acid sequence of farnesyldiphosphate synthase derived from <u>Bacillus stearothermophilus</u> can be used regardless of its origin. As such gene sources, for example, <u>Bacillus stearothermophilus</u>, <u>E. coli</u>, yeast, humans, rats etc. can be used.

The gene to be mutated is an RNA or DNA coding for a farnesyldiphosphate synthase and sensitive to treatment with a mutagen, and DNA is preferably used for to ease of handling, and especially a single-stranded DNA is preferred due to its high mutation ratio.

A single-stranded DNA can be easily prepared according to a conventional Procedure for preparing a single-stranded DNA, for example, by inserting a double-stranded DNA into a phage, introducing the phage into <u>E. coli</u> cells, culturing the <u>E. coli</u> cells and recovering the phage from the resulting lysate solution; or by introducing a desired double-stranded DNA into host cells, infecting the host cells with helper phage, culturing the host cells and recovering the phage from the resulting lysate solution.

Mutation of a gene can be carried out according to a conventional procedure for artificially mutating a gene. The mutation methods can be a physical method such as irradiation with X-rays, ultraviolet rays, etc., a chemical method such as treatment with a mutagen, a method of cis incorporation by DNA polymerase, a method using synthetic oligonucleotides etc. A chemical method is preferable for ease of operation and a high mutation ratio. As a mutagen, a nitrite, such as sodium nitrite, or the like can be used. To mutate a single-stranded DNA, a nitrite is preferable. Mutagenesis is preferably carried out at a nitrite concentration of 0.01 to 2M, for example, at about 0.1 to 1M, at a temperature of 20 to 30°C, for 10 to 120 minutes.

To select a gene coding for a protein having geranylgeranyldiphosphate synthase activity from the genes subjected to the mutagenesis, the gene subjected to the mutagenesis is inserted in an expression vector, the vector is introduced into host cells, the enzyme is expressed, and the expression product is tested for geranylgeranyldiphosphate synthase

activity. Geranylgeranyldiphosphate is converted to phytoene by a phytoene synthase, and the phytoene is converted to lycopene having red color by a phytoene desaturase.

Accordingly, for example, a gene coding for a phytoene synthase and a gene coding for phytoene desaturase are inserted into an expression vector, the vector is introduced into host cells such as <u>E. coli</u> cells, and further an expression plasmid comprising a DNA to be tested is introduced into said host cells, and the double transformed host cells are cultured. If the gene to be tested encodes a geranylgeranyldiphosphate synthase, and the geranylgeranyldiphosphate produced by the gene expression is converted to phytoene and further to lycopene, the cells are red-colored. Accordingly, a desired gene can be selected very easily and efficiently by selecting a red-colored colony.

The present invention provides a protein having geranylgeranyldiphosphate synthase activity, i.e., a geranylgeranyldiphosphate synthase, having an amino acid sequence modified from a native amino acid sequence of a farnesyldiphosphate synthase. Here, the modification of an amino acid sequence means replacement of one or a few amino acids with other amino acids, deletion of one or a few amino acids or addition of one or a few amino acids, or a combination of these modifications. The amino acid replacement is especially preferable. Regarding the number of amino acids to be modified, "a few amino acids" means usually about 15 amino acids, preferably about 10 amino acids, and more preferably about 5 amino acids. Namely, according to the present invention, the number of mutated amino acids is about 1 to 15, preferably about 1 to 10, and more preferably 1 to 5.

To determine the positions of modified amino acids, after the mutagenesis and the selection of a gene coding for a geranylgeranyldiphosphate synthase, a nucleotide sequence of the selected gene is determined, and an amino acid sequence is predicted from the determined nucleotide sequence, the predicted amino acid sequence of the modified enzyme is composed with the corresponding native amino acid sequence. Amino acid sequences thus determined of the modified enzymes are shown in Fig. 4.

In Fig. 4, the row indicated by the symbol W.T shows, by the one-letter expression, a native amino acid sequence of farnesyldiphosphate synthase of <u>Bacillus stearothermophilus</u> origin, and the rows Nos. 1 to 4 show representative amino acid sequences which acquired geranylgeranyldiphosphate synthase activity by amino acid replacement in the amino acid sequence of the farnesyldiphosphate synthase, wherein only the amino acids different from the corresponding amino acids in the native amino acid sequence of the farnesyldiphosphate synthase shown in the line T.W are indicated by the one-letter expression of amino acid.

The modified enzyme No. 1 has two mutations, i.e., the 81st position (Tyr \rightarrow His) and 275th position (Leu \rightarrow Ser); the modified enzyme No. 2 has two mutations, i.e., 34th position (Leu-Val) and 59th position (Arg \rightarrow Gln); the modified enzyme No. 3 has two mutations, i.e., 157th position (Val \rightarrow Ala) and 182nd position (His \rightarrow Tyr); and the modified enzyme No. 4 has three mutations, i.e., 81st position (Tyr \rightarrow His), 238th position (Pro \rightarrow Arg) and 265th position (Ala \rightarrow Thr). The amino acid sequences No. 1 to 4 of the above-mentioned modified enzymes and nucleotide sequences coding therefor are shown in SEQ ID NO: 1 to 4, and the native amino acid sequence and a nucleotide sequence coding therefor is shown in SEQ ID NO: 5.

30

35

40

In the present invention, the amino acid sequence farnesyldiphosphate synthase of <u>Bacillus stearothermophilus</u> origin was used as a specific example. However, as shown in Figs. 2 and 3, farnesyldiphosphate synthases have high homology among a wide spectrum of species covering those derived from the eukaryotes including humans and those derived from prokaryotes including bacteria. Therefore, the present invention can be applied to enzymes derived from various species to obtain novel geranylgeranyldiphosphate synthase.

As shown in Fig. 4, amino acid modification such as replacement occurs on the 34th, 59th, 81st, 157th, 182nd, 239th, 265th, and/or 275th positions of farnesyldiphosphate of <u>Bacillus stearothermophilus</u>. For enzymes from other species, it is expected that replacement at positions corresponding to the above-mentioned positions of the farnesyldiphosphate synthase of <u>Bacillus stearothermophilus</u> origin provides similar effects as that for the modified enzyme derived from <u>Bacillus stearothermophilus</u>. Therefore, the present invention can be applied to any farnesyldiphosphate synthases.

The present invention also relates to genes coding for the various geranylgeranyldiphosphate synthases derived from a farnesyldiphosphate synthase. These genes can be obtained by mutation of a gene coding for a corresponding native amino acid sequence. In addition, once the position of mutated amino acid is determined, a gene coding for the modified enzyme can be obtained by site-specific mutagenesis using a mutagenic primer. In addition, once an entire amino acid sequence is determined, a DNA coding for the amino acid sequence can be chemically synthesized according to a conventional procedure.

Genes coding for farnesyldiphosphate synthases used as starting materials to obtain the present genes have been cloned from various organisms, and therefore they can be used. For example, a gene of <u>Bacillus stearothermophilus</u> origin is described in J. Biochem. <u>113</u>, 355 - 363 (1993), a gene of <u>E. coli</u> origin is described in J. Biochem. <u>108</u>, 995 - 1000 (1990), a gene of yeast origin is described in J.B.C. <u>264</u>, 19176 - 19184 (1989), a gene of rat origin is described in Mol. Cell. Biol. <u>7</u>, 3138 - 3146 (1987), and a gene of human origin is described in J.B.C. <u>265</u>, 4607 - 4614 (1990).

The present invention further provides recombinant vectors, especially expression vectors, comprising the abovementioned gene (DNA), recombinant host transformed with said vector, and a process for production of said enzyme using said recombinant host.

As an example, where <u>E. coli</u> is used as a host, it is known that there are gene expression control mechanisms which regulate transcription of DNA to mRNA, translation of mRNA to protein etc.

As promoter sequences which control the synthesis of mRNA, naturally occurring sequences such as lac, trp, bla, lpp, PL, PR, tet, T3, T7 et al., as well as mutants thereof, such as lacUV5, sequences prepared by fusing naturally occurring promoter sequences, such as tac, tra, etc. are known, and they can be used in the present invention.

As sequences which control the ability to synthesize a protein from mRNA, it is known that a ribosome-binding site (GAGG and similar sequence) and the distance between the ribosome-binding site and the start codon ATG are important. In addition, it is known that a terminator which directs the termination of transcription at the 3'-end (for example, a vector comprising rrnBT1T2 is commercially available from Pharmacia) influences the efficiency of protein synthesis in a recombinant host.

As starting vectors to prepare recombinant vectors of the present invention, those commercially available can be used. Alternatively, various vectors derivatized according to a particular purpose can be used. For example, pBR322, pBR327, pKK223-2, pKK233-2, pTrc99A etc. containing a replicon derived from pMB1; pUC18, pUC19, pUC118, pUC119, pTV118N, pTV119N, pHSG298, pHSG396 etc., which have been modified to increase copy number; pACYC177, pACYC184 etc. containing a replicon derived from p15A; as well as plasmids derived from pSC101, C01E1, R1 or F-factor, may be mentioned.

Further, in addition to plasmids, viral vectors such as λ phage, M13 phage etc., and transposones can be used for introduction of a gene. These vectors are described in Molecular cloning (J. Sambrook, E.F. Fritsch, J. Maniatis, Cold Spring Harbor Laboratory Press); Cloning vector (P.H. Pouwels, B.E. Enger-Valk, W.J. Brammer, Elsevier); and catalogs of manufacturers of vectors.

Especially preferable is pTrc99 (commercially available for Pharmacia) which has an ampicillin resistance gene as a selective maker, Ptrc and lacl^q as a promoter and control gene, an AGGA sequence as a ribosome-binding site and rrnBT1T2 as a terminator, and therefore has a function to control an expression of a geranylgeranyldiphosphate synthese.

Introduction of a DNA coding for geranylgeranyldiphosphate synthase and if necessary DNA fragments having a function to control the expression of said gene into the above-mentioned vectors can be carried out using appropriate restriction enzymes and ligases according to a conventional procedure.

Such a recombinant vector can be used to transform a microorganism such as <u>Escherichia coli, Bacillus</u> etc. Transformation can be carried out according to a conventional procedure, for example by the CaCl₂ method, protoplast method etc. described, for example, in Molecular cloning (J. Sambrook, E.F. Fritsch, T. Maniatis, Cold Spring Harbor Laboratory Press), DNA cloning Vol. I to III (D.M. Glover, IRLPRESS).

Although methods for expression of the present gene in <u>E. coli</u> was described in detail, according to the present invention, a DNA coding for a geranylgeranyldiphosphate synthase is inserted into a conventional expression vector according to a conventional procedure, and the vector is used to transform a host, for example, prokaryotic cells such as various bacterial cells, lower eukaryotic cells for example single cell hosts, for example, yeast cells, or higher eukaryotic cells such as silk-worm. After transformation, the transformant is cultured to produce a geranylgeranyldiphosphate synthase, according to a conventional process.

When a transformant host such as <u>E. coli</u> is cultured, geranylgeranyldiphosphate synthase is intracellularly accumulated. To recover the geranylgeranyldiphosphate from the cultured host cells, the cells are treated physiologically or chemically, for example, with a cell lysating agent to lyze the cells. The cell debris is removed, and the supernatant is subjected to an isolation process conventional for purification of enzymes. The above-mentioned cell-lysing enzyme is preferably lysozyme, and the physical treatment is preferably treatment with ultrasonic radiation. When the supernatant is heated to a temperature of about 55°C, proteins intrinsic to <u>E. coli</u> are insolubilized and removed as an insoluble precipitate. To purify the enzyme, gel-filtration chromatography, ion exchange chromatography, hydrophobic chromatography, reversed chromatography, and affinity chromatography can be used alone or in combination. During the purification and isolation steps, the desired enzyme can be stabilized by addition of a reducing agent such as dithiothreitol, protecting agent against proteases such as PMSF, BSA etc., metal ions such as magnesium, alone or in combination.

The present invention further provides a process for production of geranylgeranyldiphosphate or geranylgeranyol. In this process, isopentenyldiphosphate, dimethylallyldiphosphate, geranyldiphosphate, farnesyldiphosphate may be used as substrates.

EXAMPLES

50

5

25

30

Next, the present invention is explained in more detail by means of examples, though the present invention is not limited thereto.

Example 1. Construction of mutated genes (Fig. 5)

The translation start codon in plasmid pFE15 (Japanese Unexamined Patent Publication (Kokai) No. 5-219761) containing a gene coding for farnesyldiphosphate synthase of <u>Bacillus stearothermophilus</u> origin was changed to ATG to obtain plasmid pEX11 (J. Biochem. 113, 355 - 363 (1993)) for overexpression of farnesyldiphosphate synthase, and the plasmid pEX11 was used in the following Examples. The mutation was carried out according to M. Myers et al. (Science, 229, 242 - 247 (1985)).

First, a farnesyldiphosphate synthase gene present in Ncol-HindIII fragment in pEX11 was removed, and inserted it into plasmid pTV118N (available from Takara Shuzo, Japan) to construct a plasmid, which was then introduced into E. coli cells. The transformed E. coli cells were cultured. With infection of a helper phage M13K07 (available from Takara Shuzo), pTV118N is converted to a single-stranded DNA and preferentially incorporated in phage particles and liberated out of cells. The culture was centrifuged to obtain a supernatant, from which the single-stranded DNA was recovered.

The single-stranded DNA thus recovered was subjected to mutation with sodium nitrite (concentration 1M or 0.2M) to introduce random mutation into the single-stranded DNA, which was then restored to a double-stranded DNA using AMV reverse-transcriptase XL (E.C.2.7.7.7). This farnesyldiphosphate synthase gene fragment was introduced into pTrc99A (available for Pharmacia) and pTV118N, and resulting recombinant plasmids were used to transform <u>E. coli</u> into which a phytoene synthase gene and phytoene desaturase gene had been previously introduced, and red colonies were selected. The principle of the selection is as follows.

The following screening method follows Ohnuma et al. (J. Biol. Chem., <u>269</u>, 14792 - 14797 (1994)). <u>E. coli</u> harboring a plasmid pACYC-IB, into which crtB (phytoene synthase gene) and crtI (phytoene desaturase gene) of a phytopathogen <u>Erwinia uredovora</u> origin had been introduced, was transformed with the mutant plasmid. Note that at present it is believed that <u>E. coli</u> does not have a geranylgeranyldiphosphate synthase. If the mutant plasmid encodes geranylgeranyldiphosphate synthase activity, lycopene having red color is produced in <u>E. coli</u> cells by pACYC-IB resulting in formation of red-colored colonies. However, if the mutant plasmid does not encode geranylgeranyldiphosphate synthase activity, colonies are color-less. In this way, geranylgeranyldiphosphate synthase activity was easily detected by visual observation.

As a result of transformation of the <u>E. coli</u> cells with the mutant plasmid, red colonies were detected. The ratio of positive clones was 1.32×10^{-3} (10 colonies per 7,600 colonies) when the mutation was carried out using 1M NaNO₂, while the ratio of positive clones was 5.98×10^{-5} (one colony per 16,720 colonies) when the mutation was carried out using 0.2M NaNO₂, revealing that the higher the concentration of NaNO₂, the higher the positive ratio.

Among the positive colonies, four colonies were selected, and a nucleotide sequence of an enzyme-coding region in the plasmid was determined, and an amino acid sequence encoded by the nucleotide sequence was determined, for each positive clone. The result is shown in SEQ ID NOs: 1 to 4. In addition, these amino acid sequences were compared with the native amino acid sequence, and positions of the mutation are shown in Fig. 4.

Four mutated enzymes encoded by four mutant genes were further characterized.

Example 2. Production of mutated enzymes

E. coli transformed with the mutant plasmid was cultured in LB medium at 37°C overnight. The culture was centrifuged at 3,000 \times G, at 4°C for 5 minutes to collect cells, which were then suspended in a buffer for sonication (50 mM Tris-HCl (pH 7.0), 10 mM 2-mercaptoethanol, 1 mM EDTA). The suspension was subjected to ultrasonic waves to disrupt the cells. The sonicate was centrifuged at 5,000 \times g, at a temperature of 4°C for 20 minutes, to obtain a supernatant, which was then heated at 55°C for one hour to inactivate enzymes intrinsic to E. coli to obtain a crude enzyme extract.

55

50

40

To test the enzymatic activity of each mutant enzyme, reactions were carried out in the following reaction mixture.

Table 1

[1-14C]IPP (1 Ci/mol))	25 nmol
Allyl substrate (DMAPP, GPP, FPP)	25 nmol
MgCl ₂	5 μmol
NH ₄ Cl	50 μmol
2-Mercaptoethanol	50 μmol
Tris-HCl buffer (pH 8.5)	50 µmol
Sample to be tested	proper quantity
Total	1 ml
Note:	

Note:

DMAPP: Dimethylallyldiphosphate

GPP: Geranyldiphosphate FPP: Farnesyldiphosphate

The reaction mixture was incubated at 55°C for 30 minutes, and the product was extracted with water-saturated 1-butanol, and radioactivity of the extract was counted by a liquid scintillation counter. In addition, the extract (butanol layer) was treated with an acid phosphatase and extracted with pentane. The extract was analyzed by TLC. The TLC analysis showed that the use of dimethylallyldiphosphate and geranyldiphosphate as an allyl substrate provides similar TLC patterns. Note that since the amount of each sample was adjusted so that the radioactivity is approximately same between the samples, the density of the band does not indicate specific activity.

The modified enzymes Nos. 1 and 4 produced an amount of geranylgeranyldiphosphate more than that of farnesyldiphosphate, and therefore it is considered that the modified enzymes Nos. 1 and 4 are suitable for the production of geranylgeranyldiphosphate. On the other hand, the modified enzymes No. 2 and No. 3 provided a small amount of geranylgeranyldiphosphate.

Where (all-E)-farnesyldiphosphate was used as a substrate (primer), (all-E)-geranylgeranyldiphosphate was formed. The results are shown in Figs. 6 to 9.

Specific activity and ratio of product (GGOH/FOH) are shown in Table 2.

Table 2

		Specific activity* (nmol/min/mg protein)	Ratio of product (GGPP/FPP)
W	ild type	286	0
No. 1	pTV118N	0.293	18.4
	pTrc99A	0.253	6.28
No. 2	pTV118N	110	2.95 × 10 ⁻²
1.	pTrc99A	83	2.54 × 10 ⁻²
No. 3	pTV118N	143	1.65 × 10 ⁻¹
	pTrc99A	19.7	. 1.73 × 10 ⁻¹
No. 4	pTV118N	0.262	15.5
	pTrc99A	0.271	8.28

*DMAPP was used as substrate

50

5

10

15

20

30

SEQUENCE LISTING

55

5	SE	Q II	O NO):	1												
	SE	QUE	NCE	LEN	GTH :	: 8	94										
	SE	QUE	NCE	TYP	E:	Nuc	lei	c ac	cid							•	
10	STI	RANI	ONES	s:	Dou	ıble								•			
	TOI	POLO	GY:	L.	inea	ır											
	MOI	ECI	JLAR	TΥ	PE:												
15		JRCE			illu		+021	coth	Orm	onh	: 1.,,						
										_				_			
			TER							t Di	NA C	odı	ng i	tor			
20	far	nes	yld	ipho	osph	ate	syr	ntha	se								
	SEÇ	UEN	ICE														
25																GCG	48
25	Met	Ala	Gln	Leu	Ser	Va1	Glu	Gln	Phe	Leu	Asn	Glu	Gln	Lys	G1n	Ala	
					5					10					15		
30																GCG	96
	vai	GIU	Inr	A18 20	Leu	ser	Arg	Tyr	11e 25	Glu	Arg	Leu	Glu	Gly 30	Pro	Ala	
	AAG	CTG	AAA		GCG	ATG	GCG	TAC		ፐፐር	GAG	GCC	ccc		AAA	CGA	144
35													Gly				177
			35					40		٠		•	45	·		Ū	
	ATC	CGT	CCG	TTG	CTG	CTT	CTG	TCC	ACC	GTT	CGG	GCG	СТС	GGA	AAA	GAC	192
40	Ile	Arg	Pro	Leu	Leu	Leu	Leu	Ser	Thr	Val	Arg	Ala	Leu	Gly	Lys	Asp	
•		50					55					60					
4 5													ATG				240
*5		Ala	Val	G1y	Leu		Va1	Ala	Cys	Ala	Ile	Glu	Met	Ile	His	Thr	
	65	TCT.	mm.c	4.50	0 4 m	70	0 4 m				75					80	
50													AAC				288
	****	Jel	Deu	116	85	vah	vsh	Leu	LIO	90	ue C	wab	Asn	Asp	Asp 95	ren	
										70					9.5		

	CGG	CGC	GGC	AAG	CCC	ACG	AAC	CAT	AAA 1	GTC	TTC	GGC	GAG	GCG	ATC	GCC	336
_	Arg	Arg	g Gly	Lys	Pro	Thr	Asn	His	Lys	Val	Phe	Gly	Glu	Ala	Met	Ala	
5				100)				105					110		•	
	ATC	TTG	GCG	GGG	GAC	GGG	TTG	TTG	ACG	TAC	GCG	TTT	CAA	TTG	ATC	ACC	384
	Ile	Leu	Ala	Gly	Asp	Gly	Leu	Leu	Thr	Tyr	Ala	Phe	Gln	Leu	lle	Thr	
10			115	•				120					125				
	GAA	ATC	GAC	GAT	GAG	CGC	ATC	CCT	CCT	TCC	GTC	CGG	CTT	CGG	CTC	ATC	432
4-	Glu	Ile	Asp	Asp	Glu	Arg	Ile	Pro	Pro	Ser	Val	Arg	Leu	Arg	Leu	Ile	
15		130					135					140					
	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GTC	GCC	GGT	CAG	480
00	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	G1y	Met	Val	Ala	Gly	G1n	
20	145					150					155					160	
	GCA	GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	528
<i>25</i>	Ala	Ala	Asp	Met	Glu	Gly	Glu	G1y	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu	
25					165	`				170					175		
	GAA	TAC	ATT	CAT	CGG	CAT	AAA	ACC	GGG	AAA	ATG	CTG	CAA	TAC	AGC	GTG	576
30	G1u	Tyr	Ile	His	Arg	His	Lys	Thr	Gly	Lys	Met	Leu	Gln	Tyr	Ser	Val	
00				180					185					190			
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
35	His	Ala	Gly	Ala	Leu	Ile	Gly	Gly	Ala	Asp	Ala	Arg	Gl'n	Thr	Arg	Glu	
			195					200					205				
	CTT	GAC	GAA	TTC	GCC	GCC	CAT	CTA	GGC	CTT	GCC	TTT	CAA	ATT	CGC	GAT	672
40	Leu	Asp	Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	Gln	Ile	Arg	Asp	
		210					215					220			9		
	GAT	ATT	CTC	GAT	ATT	GAA	GGG	GCA	GAA	GAA	AAA	ATC	GGC	AAG	CCG	GTC	720
45	Asp	Ile	Leu	Asp	Ile	G1u	Gly	Ala	Glu	Glu	Lys	Ile	Gly	Lys	Pro	Va1	
	225					230					235					240	
	GGC	AGC	GAC	CAA	AGC	AAC	AAC	AAA	GCG	ACG	TAT	CCA	GCG	TTG	CTG	TCG	768
50	Gly	Ser	Asp	Gln	Ser	Asn .	Asn	Lys	Ala	Thr	Tyr	Pro	Ala :	Leu	Leu	Ser	
-				•	245					250	٠				255		

	CTT	. ecc	GGC	GCG	AAG	GAA	AAG	TTG	GCG	TTC	CAT	ATC	GAG	GCG	GCG	CAG	816
5	Leu	Ala	Gly	Ala	Lys	Glu	Lys	Leu	Ala	Phe	His	Ile	Glu	Ala	Ala	Gln	
J				260					265					270			
	CGC	CAT	TCA	CGG	AAC	GCC	GAC	GTT	GAC	GGC	GCC	GCG	CTC	GCC	TAT	ATT	864
10	Arg	His	Ser	Arg	Asn	Ala	Asp	Val	Asp	Gly	Ala	Ala	Leu	Ala	Tyr	Ile	
	•		275					280					285			•	
			CTG														894
15	Cys		Leu	Val	Ala	Ala	_	Asp	His	***				•			
	ana	290		•			295										
		_	NO:														
20	SEÇ)UEN	CE I	LENG	TH:	89	94										
	SEÇ	UEN	CE 1	YPE	: 1	Nuc]	leic	ac	id								
	STF	RAND	NESS	:	Dou	ble											
25	TOF	oro	Y:	Lin	ear												
	MOL	ECU	LAR	TYP	E :		*										
30	sou	IRCE	: E	Baci	llus	s st	ear	othe	ermo	phi	lus						
	СНА	RAC	TERI	STI	C:	Mut	ant	(2)) of	DN	A c	odin	ıg f	or			
	far	nes	yldi	.pho	spha	ate	sys	thas	se								
35	SEQ	UEN	CE														
	ATG	GCG	CAG	CTT	TCA	GTT	GAA	CAG	TTT	CTC	AAC	GAG	CAA	AAA	CAG	GCG	48
40			Gln														
••					5					10				•	15		
	GTG	GAA	ACA	GCG	CTC	TCC	CGT	TAT	ATA	GAG	CGC	TTA	GAA	GGG	CCG	GCG	96
45	Val	Glu	Thr	Ala	Leu	Ser	Arg	Tyr	Ile	Glu	Arg	Leu	Glu	Gly	Pro	Ala	
				20					25					30			
	AAG	GTG	AAA	AAG	GCG	ATG	GCG	TAC	TCA	TTG	GAG	GCC	GGC	GGC	AAA	CGA	144
50	Lys	Val	Lys	Lys	Ala	Met	Ala	Tyr	Ser	Leu	Glu	Ala	Gly	Gly	Lys	Arg	
			35					40					45				

	ATC	CGT	CCG	TTG	CTG	CTT	CTG	TCC	ACC	GTT	CAG	GCG	CTC	GGC	AAA	GAC	192
5	Ile	Arg	Pro	Leu	Leu	Leu	Leu	Ser	Thr	Val	Gln	Ala	Leu	Gly	Lys	Asp	
		50					55					60					
	CCG	GCG	GTC	GGA	TTG	CCC	GTC	GCC	TGC	GCG	ATT	GAA	ATG	ATC	CAT	ACG	240
10	Pro	Ala	Val	Gly	Leu	Pro	Val	Ala	Cys	Ala	.Ile	Glu	Met	Ile	His	Thr	
	65					70					75					80	
	TAC	TCT	TTG	ATC	CAT	GAT	GAT	TTG	CCG	AGC	ATG	GAC	AAC	GAT	GAT	TTG	288
1 5	Tyr	Ser	Leu	Ile	His	Asp	Asp	Leu	Pro	Ser	Met	Asp	Asn	Asp	Asp	Leu	
,,,					85					90					95		
	CGG	CGC	GGC	AAG	ÇCG	ACG	AAC	CAT	AAA	GTG	TTC	GGC	GAG	GCG	ATG	GCC	336
20	Arg	Arg	Gly	Lys	Pro	Thr	Asn	His	Lys	Val	Phe	Gly	Glu	Ala	Met	Ala	
			•	100					105					110			
	ATC	TTG	GCG	GGG	GAC	GGG	TTG	TTG	ACG	TAC	GCG	TTT	CAA	TTG	ATC	ACC	384
25	Ile	Leu	Ala	Gly	Asp	Gly	Leu	Leu	Thr	Týr	Ala	Phe	Gln	Leu	Ile	Thr	
		•	115					120					125				
	GAA	ATC	GAC	GAT	GAG	CGC	ATC	CCT	CCT	TCC	GTC	CGG	CTT	CGG	CTC	ATC	432
30	Glu	Ile	Asp	Asp	G1u	Arg	Ile	Pro	Pro	Ser	Val	Arg	Leu	Arg	Leu	Ile	
	•	130					135					140					
	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GTC	GCC	GGT	CAG	480
35	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	Gly	Met	Val	Ala	Gly	Gln	
	145				•	150					155			,		160	
	GCA	GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ĄCG	CTG	AÇG	CTT	TCG	GAG	CTC	528
10	Ala	Ala	Asp	Met	Glu	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu	
					165					170					175		
	GAA	TAC	ATT	CAT	CGG	CAT	AAA	ACC	GGG	AAA	ATG	CTG	CAA	TAC	AGC	GTG	576
15	Glu	Tyr	Ile	His	Arg	His	Lys	Thr	Gly	Lys	Met	Leu	Gln	Tyr	Ser	Va1	
				180					185					190			
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
io	His	Ala	Gly	Ala	Leu	Ile	Gly	Gly	Ala	Asp	Ala	Arg	G1n	Thr	Arg	Glu	
			195					200					205				

	CTT	GAC	GAA	TTC	GCC	GCC	CAT	CTA	GGC	CTT	GCC	TTT	CAA	ATT	CGC	GAT	672
_	Leu	Asp	Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	G1n	Ile	Arg	Asp	
5		210					215					220					
	GAT	ATT	CTC	GAT	ATT	GAA	GGG	GCA	GAA	GAA	AAA	ATC	GGC	AAG	CCG	GTC	720
40	Asp	Ile	Leu	Asp	Ile	Glu	Gly	Ala	Glu	Glu	Lys	Ile	Gly	Lys	Pro	Val	
10	225					230					235					240	
	GGC	AGC	GAC	CAA	AGC	AAC	AAC	AAA	GCG	ACG	TAT	CCA	GCG	TTG	CTG	TCG	768
15	Gly	Ser	Asp	Gln	Ser	Asn	Asn	Lys	Ala	Thr	Tyr	Pro	Ala	Leu	Leu	Ser	
75					245					250					255		
	CTT	GCC	GGC	GCG	AAG	GAA	AAG	TTG	GCG	TTC	CAT	ATC	GAG	GCG	GCG	CAG	816
20	Leu	Ala	Gly	Ala	Lys	Glu	Lys	Leu	Ala	Phe	His	Ile	Glu	Ala	Ala	Gln	
				260					265					270			
	CGC	CAT	TTA	CGG	AAC	GCC	GAC	GTT	GAC	GGC	GCC	GCG	CTC	GCC	TAT	ATT	864
25	Arg	His	Leu	Arg	Asn	Ala	Asp	Val	Asp	Gly	Ala	Ala	Leu	Ala	Tyr	Ile	
			275					280				•	285				
	TGC	GAA	CTG	GTC	GCC	GCC	CGC	GAC	CAT	TAA							894
30	Cys	Glu	Leu	Val	Ala	Ala	Arg	Asp	His	***							
		290					295										
	SEQ	ID	NO:	3													
35	SEQ	UENC	CE L	ENG	TH:	89	4										
	SEQ	UENC	CE T	YPE	: 1	Nucl	eic	aci	id								
40	STR	ANDI	NESS	: 1	Doul	ole											
70	TOP	OLO	<i>:</i>	Lin	ear												
	MOL	ECUI	LAR	TYP:	E:												
45	sou	RCE:	: B	aci	llus	sst	ear	othe	ermo	phi	lus						
	CHA	RACI	reri	STI	C:	Mut	ant	(3)	of	DN	A co	odin	g f	or			
50	far	nesy	/ldi	pho	spha	ate	sys	thas	se								
	SEQ	UENC	CE														

	ATG	GCG	CAG	CTT	TCA	GTT	GAA	CAG	TTT	CTC	AAC	GAG	CAA	AAA	CAG	GCG	48
5	Met	Ala	Gln	Leu	Ser	Val	Glu	Gln	Phe	Leu	Asn	Glu	Gln	Lys	Gln	Ala	
5					5					10					15		
	GTG	GAA	ACA	GCG	CTC	TCC	CGT	TAT	ATA	GAG	CGC	TTA	GAA	GGG	CCG	GCG	96
10	Val	Glu	Thr	Ala	Leu	Ser	Arg	Tyr	Ile	Glu	Arg	Leu	G1u	Gly	Pro	Ala	
70				20					25					30			
·	AAG	CTG	AAA	AAG	GCG	ATG	GCG	TAC	TCA	TTG	GAG	GCC	GGC	GGC	AAA	CGA	144
15	Lys	Leu	Lys	Lys	Ala	Met	Ala	Tyr	Ser	Leu	Glu	Ala	Gly	Gly	Lys	Arg	
			35					40					45				
	ATC	CGT	CCG	TTG	CTG	CTT	CTG	TCC	ACC	GTT	CGG	GCG	CTC	GGA	AAA	GAC	192
20	Ile	Arg	Pro	Leu	Leu	Leu	Leu	Ser	Thr	Val	Arg	Ala	Leu	Gly	Lys	Asp	
		50					55					60			,		
	CCG	GCG	GTC	GGA	TTG	CCC	GTC	GCC	TGC	GCG	ATT	GAA	ATG	ATC	CAT	ACG	240
25	Pro	Ala	Val	Gly	Leu	Pro	Val	Ala	Cys	Ala	Ile	Glu	Met	Ile	His	Thr	
	65					70					75					80.	
	TAC	TCT	TTG	ATC	CAT	GAT	GAT	TTG	CCG	AGC	ATG	GÁC	AAC	GAT	GAT	TTG	288
30	Tyr	Ser	Leu	Ile	His	Asp	Asp	Leu	Pro	Ser	Met	Asp	Asn	Asp	Asp	Leu	
					85					90					95		
	CGG	CGC	GGC	AAG	CCG	ACG	AAC	CAT	AAA	GTG	TTC	GGC	GAG	GCG	ATG	GCC	336
35	Arg	Arg	Gly	Lys	Pro	Thr	Asn	His	Lys	Val	Phe	Gly	Glu	Ala	Met	Ala	
				100					105					110			
	ATC	TTG	GCG	GGG	GAC	GGG .	TTG	TTG	ACG	TAC	GCG	TTT	CAA	TTG	ATC	ACC	384
10	Ile	Leu	Ala	Gly	Asp	Gly	Leu	Leu	Thr	Tyr	Ala	Phe	Gln	Leu	Ile	Thr	
			115					120					125				
	GAA	ATC	GAC	GAT	GAG	CGC	ATC	CCT	CCT	TCC	GTC	CGG	CTT	CGG	CTC	ATC	432
15	Glu	Ile	Asp	Asp	G1u	Arg	Ile	Pro	Pro	Ser	Val	Arg	Leu	Arg	Leu	Ile	
		130					135					140					
	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GCC	GCC	GGT	CAG	480 ⁻
50	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	Gly	Met	Ala	Ala	Gly	Gln	
	145					150		,			155					160	

	GCA	A GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	528
	Ala	Ala	Asp	Met	Glu	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu	
5					165					170)				175		
	GAA	TAC	ATT	CAT	CGG	TAT	AAA	ACC	GGG	AAA	ATG	CTG	CAA	TAC	AGC	GTG	576
	Glu	Tyr	Ile	His	Arg	Tyr	Lys	Thr	Gly	Lys	Met	Leu	Gln	Tyr	Ser	Val	
10				180					185		,			190			
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
	His	Ala	Gly	Ala	Leu	Ile	Gly	Gly	Ala	Asp	Ala	Arg	Gln	Thr	Arg	Glu	
15			195					200					205				
	CTT	GAC	GAA	TTC	ĢCC	GCC	CAT	CTA	GGC	CTT	GCC	TTT	CAA	ATT	CGC	GAT	672
	Leu	Asp	Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	Gln	Ile	Arg	Asp	
20		210					215					220					
		ATT															720
		Ile	Leu	Asp	Ile	Glu	Gly	Ala	Glu	Glu	Lys	Ile	Gly	Lys	Pro	Val	
25	225					230	•				235					240	
		AGC															768
	Gly	Ser	Asp	Gln		Asn	Asn	Lys	Ala		Tyr	Pro	Ala	Leu		Ser	
30					245					250					255		
		GCC															816
	Leu	Ala	Gly		Lys	Glu	Lys	Leu		Phe	His	Ile	Glu		Ala	Gln	
35				260					265					270			•
		CAT															864
	Arg	His		Arg	Asn	Ala	Asp		Asp	Gly	Ala	Ala		Ala	Tyr	Ile	٠
40			275	0.00				280					285				
		GAA															894
	Cys	Glu 290	Leu	vai	AIA		Arg 295	Asp	HIS	***							
45	SEO		NO.	4			293										
	•	ID															
	SEQ	UENC	E L	ENG'	TH:	89	4										
50	SEQ	UENC	E T	YPE	: N	lucl	eic	aci	.d								
	STR	ANDN	IESS	: i	Doub	le											
	тор	OLOY	7:	Line	ear												

MOLECULAR TYPE:

55

	500	,,,c	•	Duca			ccui		CIM	Opiii		•					
5	CHA	RAC	TER	ISTI	C:	Mu	tant	- (4) 0	f Di	NA c	odi	ng i	for		٠	
	far	nes	yld	ipho	sph	ate	sys	tha	se								
10	SEÇ	UEN	CE														
	ATG	GCG	CAG	CTT	TCA	GTT	GAA	CAG	TTT	CTC	AAC	GAG	CAA	AAA	CAG	GCG	48
	Met	Ala	Gln	Leu	Ser	Val	Glu	Gln	Phe	Leu	Asn	Glu	Gln	Lys	Gln	Ala	
15					5					10					15		
	GTG	GAA	ACA	GCG	CTC	TCC	CGT	TAT	ATA	GAG	CGC	TTA	GAA	GGG	CCG	GCG	96
	Val	Glu	Thr		Leu	Ser	Arg	Tyr		Glu	Arg	Leu	Glu	_	Pro	Ala	
20				20			200		25			222	000	30		204	• • •
															AAA		144
•	Lys	Leu	35	Буб	VIA	Met	AIA	40	261	Leu	GIU	NIA	45	GIY	Lys	VIR	
25	ATC	CGT		TTG	CTG	CTT	CTG		ACC	GTT	CGG	GCG		GGC	AAA	GAC	192
															Lys		
		50					55					60		·	-	-	
30	CCG	GCG	GTC.	GGA	TTG	ссс	GTC	GCC	TGC	GCG	ATT	GAA	ATG	ATC	CAT	ACG	240
	Pro	Ala	Val	Gly	Leu	Pro	Val	Ala	Cys	Ala	Ile	Glu	Met	Ile	His	Thr	
	65					_. 70					75					80	
35	CAC	TCT	TTG	ATC	CAT	GAT	GAT	TTG	CCG	AGC	ATG	GẠC	AAC	GAT	GAT	TTG	288
	His	Ser	Leu	Ile	His	Asp	Asp	Leu	Pro	Ser	Met	Asp	Asn	Asp	Asp	Leu	
					85					90					95		
40	•														ATG		336
	Arg	Arg	GIY	•	Pro	Tnr	ASN	HIS	_	val	Pne	Gly	GIU		Met	Ala	
	ΔТС	ጥ ጥር	GCG	100	GAC	GGG	ጥጥር	ጥ ተር	105	ፐልር	ccc	ጥጥጥ	CAA	110 TTG	ATC	ACC	384
45									,						Ile		304
			115	,		,		120		-,-			125				
	GAA	ATC		GAT	GAG	CGC	ATC		CCT	тсс	GTC	CGG		CGG	СТС	ATC	432
50	Glu	Ile	Asp	Asp	Glu	Arg	Ile	Pro	Pro	Ser	Val	Arg	Leu	Arg	Leu	Ile	
		130					135					140					

	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GTC	GCC	GGI	CAG	480
	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	Gly	Met	Val	Ala	Gly	Gln	
5	145					150					155					160	
	GCA	GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	528
	Ala	Ala	Asp	Met	Glu	Gly	G1u	Gly	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu	
10					165					170					175		
	GAA	TAC	ATT	CAT	CGG	CAT	AAA	ACC	GGG	ÄAA	ATG	CTG	CAA	TAC	AGC	GTG	576
	Glu	Tyr	Ile	His	Arg	His	Lys	Thr	Gly	Lys	Met	Leu	Gln	Tyr	Ser	Val	
15				180					185					190			
	CAC	GCC	GGC	GCC	ŢTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
	His	Ala	Gly	Ala	Leu	Ile	Gly	Gly	Ala	Asp	Ala	Arg	Gln	Thr	Arg	Glu	
20			195					200					205				
	CTT	GAC	GAA	TTC	GCC	GCC	CAT	CTA	GGC	CTT	GCC	TTT	CAA	ATT	CGC	GAT	672
	Leu	Asp	Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	Gln	Ile	Arg	Asp	
25		210					215					220					
	GAT	ATT	CTC	GAT	ATT	GAA	GGG	GCA	GAA	GAA	AAA	ATC	GGC	AAG	CGG	GTC	720
	Asp	Ile	Leu	Asp	Ile	Glu	Gly	Ala	Glu	Glu	Lys	Ile	Gly	Lys	Arg	Val	
30	225					230					235					240	
	GGC	AGC	GAC	CAA	AGC	AAC	AAC	AAA	GCG	ACG	TAT	CCA	GCG	TTG	CTG	TCG	768
	Gly	Ser	Asp	Gln	Ser	Asn	Asn	Lys	Ala	Thr	Tyr	Pro	Ala	Leu	Leu	Ser	
35					245					250					255		
	CTT	GCC	GGC	GCG	AAG	GAA	AAG	TTG	ACG	TTC	CAT	AŢC	GAG	GCG	GCG	CAG	816
	Leu	Ala	Gly	Ala	Lys	Glu	Lys	Leu	Thr	Phe	His	Ile	Glu	Ala	Ala	Gln	
40				260					265					270			
	CGC	CAT	TTA	CGG	AAC	GCC	GAC	GTT	GAC	GGC	GCC	GCG	CTC	GCC	TAT	ATT	864
	Arg	His	Leu	Arg	Asn	Ala	Asp	Val	Asp	Gly	Ala	Ala	Leu	Ala	Tyr	Ile	
45			275					280					285				
	TGC	GAA	CTG	GTC	GCC	GCC	CGC	GAC	CAT	TAA							894
	Cys	Glu	Leu	Val	Ala	Ala	Arg	Asp	His	***							
5 0		290					295								•		
50	SEQ	ID	NO:	5													
	SEQ	UEN	CE L	ENG	TH:	89	4										

	35,	2061	·CE	IIF	•	MUC	Ter	- ac	.14							-	
5	STI	RANE	NES	s:	Dou	ble											
	TOI	OLC	Y:	Liı	near	•						•					
	MOI	LECU	LAR	TYI	PE:												
10	sot	JRCE	:	Baci	illu	s s	teai	coth	erm	oph:	ilus	3					
•	CHA	ARAC	TER	IST	c:	DN.	A co	odin	g f	or 1	nati	.ve	farı	nesy	ldi	phos	phate
15	syr	tha	.se												-		
	SEÇ	UEN	CE														
	ATG	GCG	CAG	CTT	TCA	GTT	GAA	CAG	TTT	CTC	AAC	GAG	CAA	AAA	CAG	GCG	48
20	Met	Ala	Gln	Leu	Ser	Val	Glu	Gln	Phe	Leu	Asn	Glu	Gln	Lys	Gln	Ala	
					5					10					15		
	GTG	GAA	ACA	GCG	CTC	TCC	CGT	TAT	ATA	GAG	CGC	TTA	GAA	GGG	CCG	GCG	96
25	Val	Glu	Thr	Ala	Leu	Ser	Arg	Tyr	Ile	G1u	Arg	Leu	Glu	Gly	Pro	Ala	
				20					25	•				30			
																CGA	144
30	Lys	Lys		Lys	Ala	Met	Ala	•	Ser	Leu	Glu	Ala	•	Gly	Lys	Arg	
	4 m.c	005	35		6 m6			40					45				
							CTG										192
35		50	PIO	Leu	Leu	Leu	Leu 55	ser	inr	vaı	GIN	60	Leu	GIY	Lys	Asp	
	CCG		GTC	GGA	TTG	ccc	GTC	GCC	TGC	929	АТТ	-	ATG	ATC	CAT	ACG	240
\$ 0							Val										240
	65			·		70			•		75					80	
	TAC	тст	TTG	ATC	CAT	GAT	GAT	TTG	CCG	AGC	ATG	GAC	AAC	GAT	GAT	TTG	288
15	Tyr	Ser	Leu	Ile	His	Asp	Asp	Leu	Pro	Ser	Met	Asp	Asn	Asp	Asp	Leu	
					85					90					95		
	CGG ·	CGC	GGC	AAG	CCG	ACG	AAC	CAT	AAA	GTG	TTC	GGC	GAG	GCG	ATG	GCC	336
50	Arg	Arg	Gly	Lys	Pro	Thr	Asn	His	Lys	Va1	Phe	Gly	G1u	Ala	Met	Ala	
				100					105					110			

	ATC	TTC	GCG	GGG	GAC	GGG	TTG	TTG	ACG	TAC	GCG	TT	CAA	TTC	ATO	C ACC	. 384
	Ile	Leu	Ala	Gly	Asp	G1y	Leu	Leu	Thr	Tyr	Ala	Phe	Glr	Leu	ı Ile	Thr	
5			115	i				120					125	i			
	GAA	ATC	GAC	GAT	GAG	CGC	ATC	CCT	CCT	TCC	GTC	CGG	CTI	CGG	CTC	ATC	432
	Glu	Ile	Asp	Asp	Glu	Arg	Ile	Pro	Pro	Ser	Val	Arg	Leu	Arg	Leu	Ile	
10		130					135					140	1				
	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GTC	GCC	GGT	CAG	480
	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	G1y	Met	Val	Ala	Gly	Gln	
15	145					150					155					160	
	GCA	GCC	GAT	ATG	ĢAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	528
	Ala	Ala	Asp	Met	Glu	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Leu	Ser	G1u	Leu	
20					165					170					175		
	GAA	TAC	ATT	CAT	CGG	CAT	AAA	ACC	GGG	AAA	ATG	CTG	CAA	TAC	AGC	GTG	576
	Glu	Tyr	Ile	His	Arg	His	Lys	Thr	Gly	Lys	Met	Leu	G1n	Tyr	Ser	Va1	
25				180					185					190			
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
	His	Ala	Gly	Ala	Leu	Ile	Gly	G1y	Ala	Asp	Ala	Arg	Gln	Thr	Arg	Glu	•
30			195					200					205				
30	•		GAA														672
	Leu		Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	Gln	Ile	Arg	Asp	
		210			:		215					220					
35			CTC														720
		lle	Leu	Asp	Ile		Gly	Ala	Glu	Glu	•	Ile	Gly	Lys	Pro		
	225					230					235					240	
40	•		GAC														768
	GIY	ser	Asp	GIN		ASN	Asn	Lys	Ala		Tyr	Pro	ALA	Leu		Ser	
	ር ጥጥ	ccc	000	000	245	C4.4	446	mmc.	CCC	250	C 4 m	4.50		000	255	0.4.0	
4 5			GGC														816· 〔
	Leu	WIS	Gly		rys	GIU	Lys			Pne	HIS	TIE	GIU		Ala	GIn	
	ccc	ር ል ሞ	ጥጥል	260	A A C	ccc	CAC		265	ccc	ccc	ccc	CTC	270	T 4 T	A TO TO	0.5.4
50			TTA	•													864
	vrR	1112	Leu 275	vig	ASN -	WIS			veb	оту	WIS	WIS	285	AIA	ıyr	116	
			4/3					280					200				

TGC GAA CTG GTC GCC GCC CGC GAC CAT TAA

Cys Glu Leu Val Ala Ala Arg Asp His ***

290

295

A mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate and gene coding for said mutated enzyme, wherein the mutated enzyme is modified from a native farnesyldiphosphate synthase by mutation of a gene coding for a native farnesyldiphosphate synthase.

SEQUENCE LISTING

5	(1) GENER	RAL INFORMATION:
10 ·	(i)	APPLICANT: (A) NAME: Toyota Jidosha Kabushiki Kaisha (B) STREET: 1, Toyota-cho (C) CITY: Toyota-shi (D) STATE: Aichi (E) COUNTRY: Japan (F) POSTAL CODE (ZIP): None
15	(ii)	TITLE OF INVENTION: MUTATED FARNESYLDIPHOSPHATE SYNTHASE CAPABLE OF SYNTHESIZING GERANYLGERANYLDIPHOSPHATE AND GENE CODING THEREFOR
	(iii)	NUMBER OF SEQUENCES: 10
20	(iv)	COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
25	(v)	CURRENT APPLICATION DATA: APPLICATION NUMBER: EP 95115423.6
	(2) INFOR	MATION FOR SEQ ID NO: 1:
30	(i)	SEQUENCE CHARACTERISTICS: (A) LENGTH: 894 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear
35	(vi)	ORIGINAL SOURCE: (A) ORGANISM: Bacillus stearothermophilus
40	(ix)	FEATURE: (A) NAME/KBY: CDS (B) LOCATION: 1894 (D) OTHER INFORMATION: /function= "Mutant (1) of DNA coding for farnesyldiphosphate synthase"
45	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 1:
		AG CTT TCA GTT GAA CAG TTT CTC AAC GAG CAA AAA CAG GCG 48 In Leu Ser Val Glu Gln Phe Leu Asn Glu Gln Lys Gln Ala 5 10 15
50		CA GCG CTC TCC CGT TAT ATA GAG CGC TTA GAA GGG CCG GCG Thr Ala Leu Ser Arg Tyr Ile Glu Arg Leu Glu Gly Pro Ala 20 25 30

20

																CGA Arg	144
, 5			35					40					45	-	-		
-			Pro					Ser					Leu			GAC Asp	192
10	CCG Pro 65	GCG Ala	GTC Val	GGA Gly	TTG Leu	CCC Pro 70	Val	GCC Ala	TGC Cys	GCG Ala	ATT Ile 75	GAA Glu	ATG Met	ATC Ile	CAT His	ACG Thr 80	240
15	CAC His	TCT Ser	TTG Leu	ATC Ile	CAT His 85	GAT Asp	GAT Asp	TTG Leu	CCG Pro	AGC Ser 90	ATG Met	GAC Asp	AAC Asn	GAT Asp	GAT Asp 95	TTG Leu	288
	CGG Arg	CGC Arg	GGC Gly	AAG Lys 100	CCG Pro	ACG Thr	AAC Asn	CAT His	AAA Lys 105	GTG Val	TTC Phe	GGC Gly	GAG Glu	GCG Ala 110	ATG Met	GCC Ala	336
20		TTG Leu															384
25	GAA Glu	ATC Ile 130	GAC Asp	GAT Asp	GAG Glu	CGC Ar g	ATC Ile 135	CCT Pro	CCT Pro	TCC Ser	GTC Val	CGG Arg 140	CTT Leu	CGG Arg	CTC Leu	ATC Ile	432
30	GAA Glu 145	CGG A rg	CTG Leu	GCG Ala	AAA Lys	GCG Ala 150	GCC Ala	GGT Gly	CCG Pro	GAA Glu	GGG Gly 155	ATG Met	GTC Val	GCC Ala	GGT Gly	CAG Gln 160	480
		GCC Ala															528
35	GAA Glu	TAC Tyr	ATT Ile	CAT His 180	CGG Arg	CAT His	AAA Lys	ACC Thr	GGG Gly 185	AAA Lys	ATG Met	CTG Leu	CAA Gln	TAC Tyr 190	AGC Ser	GTG Val	576
10	CAC His	GCC Ala	GGC Gly 195	GCC Ala	TTG Leu	ATC Ile	GGC Gly	GGC Gly 200	GCT Ala	GAT Asp	GCC Ala	CGG Arg	CAA Gln 205	ACG Thr	CGG Arg	GAG Glu	624
15	CTT Leu	GAC Asp 210	GAA Glu	TTC Phe	GCC Ala	GCC Ala	CAT His 215	CTA Leu	GGC Gly	CTT Leu	GCC Ala	TTT Phe 220	CAA Gln	ATT Ile	CGC Arg	GAT Asp	672
	GAT Asp 225	ATT Ile	CTC Leu	GAT Asp	Ile	GAA Glu 230	GGG	GCA Ala	GAA Glu	GAA Glu	AAA Lys 235	ATC Ile	GGC Gly	AAG Lys	CCG Pro	GTC Val 240	720
50		AGC Ser															768

		Ala	AAG Lys		Ala			Ala		816
5		260			265			270		
			AAC Asn							864
10			GCC Ala			TAA		•		894
15										
20		•								
<i>25</i>										
30							•			
35								•		
40										

(2) INFORMATION FOR SEQ ID NO: 2:

5	-		()	A) L B) T	ENGT YPE :	CHA H: 2 ami OGY:	97 a no a	mino cid								
		(ii) MO	LECU	LB T	YPE:	pro	tein								
10		(xi) SE	OUBN	CB D	ESCR	IPTI	ON:	SEO	ID N	0: 2	:				
	Met 1		Gln			Val							Gln	Lys	Gln 15	
15	Val	Glu	Thr	Ala 20	Leu	Ser	Arg	Tyr	Ile 25	Glu	Arg	Leu	Glu	Gly 30		Ala
	Lys	Leu	Lys 35	Lys	Ala	Met	Ala	Tyr 40	Ser	Leu	Glu	Ala	Gly 45	Gly	Lys	Arg
20	Ile	Arg 50	Pro	Leu	Leu	Leu	Leu 55	Ser	Thr	Val	Arg	Ala 60	Leu	Gly	Lys	Asp
	Pro 65	Ala	Val	Gly	Leu	Pro 70	Val	Ala	Сув	Ala	Ile 75	Glu	Met	Ile	His	Thr 80
25	His	Ser	Leu	Ile	His 85	Asp	Asp	Leu	Pro	Ser 90	Met	Asp	Asn	Asp	Asp 95	Leu
	Arg	Arg	Gly	Lys 100	Pro	Thr	Asn	His	Lys 105	Val	Phe	Gly	Glu	Ala 110	Met	Ala
30	Ile	Leu	Ala 115	Gly	Asp	Gly	Leu	Leu 120	Thr	Tyr	Ala	Phe	Gln 125	Leu	Ile	Thr
35	Glu	11e 130	Asp	Asp	Glu	Arg	Ile 135	Pro	Pro	Ser	Val	Arg 140	Leu	Arg	Leu	Île
	Glu 145	Arg	Leu	Ala	Lys	Ala 150	Ala	Gly	Pro	Glu	Gly 155	Met	Val	Ala	Gly	Gln 160
40	Ala	Ala	Asp	Met	Glu 165	Gly	Glu	Gly	Lys	Thr 170	Leu	Thr	Leu	Ser	Glu 175	Leu
	Glu	Tyr	Ile	His 180	Arg	His	Lys	Thr	Gly 185	Lys	Met	Leu	Gln	Tyr 190	Ser	Val
4 5	His	Ala	Gly 195	Ala	Leu .	Ile	Gly	Gly 200	Ala	Ąsp	Ala	Arg	Gln 205	Thr	Àrg	Glu
	Leu	Asp 210	Glu	Phe	Ala	Ala	His 215	Leu	Gly	Leu	Ala	Phe 220	Gln	Ile	Arg	Asp
50	Asp 225	Ile	Leu	Asp	Ile	Glu 230	Gly	Ala	Glu	Glu	Lys 235	Ile	Gly	Lys	Pro	Val 240
	Gly	Ser	Asp	Gln	Ser 245	Asn	Asn	Lys	Ala	Thr 250	Tyr	Pro	Ala	Leu	Leu 255	Ser
<i>55</i>																

	Leu	Ala	Gly	Al a 260	Lys	Glu	Lys	Leu	Ala 265	Phe	His	Ile	Glu	Ala 270	Ala	Gln
5	Arg	His	Ser 275	Arg	Asn	Ala	Asp	Val 280	Asp	Gly	Ala	Ala	Leu 285	Ala	Tyr	Ile -
10	Cys	Glu 290	Leu	Val	Ala	Ala	Arg 295	Asp	His							
15		,														
20																
25																
30																
35																
10											٠					
15																
60																
55																

	(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:	3:							
5		(i	(A) L B) T C) S	CE C BNGT YPE: TRAN OPOL	H: 8 nuc DEDN	94 b leic ESS:	ase aci dou	pair d	s						
10		(vi			AL S RGAN			illu	s st	earo	ther	moph	ilus			
15		(ix	. (1	A) N B) L	ame/ Ocat Ther	ION:	1 ORMA	894 TION							DNA .	
		(xi) SE	QUEN	CE D	BSCR:	IPTI	ON:	SEQ	ID N	0: 3	:				
20					TCA Ser 5											48
25					CTC Leu											96
3 <i>0</i>					GCG Ala											144
					CTG Leu											192
35 -					TTG Leu											240
4 0					CAT His 85											288
4 5					CCG Pro											336
					GAC Asp											384
50					GAG Glu											432

			GCG Ala 150							480
5			GGA Gly							528
10			CAT His							576
15		Ala	ATC Ile							624
20			GCC Ala							672
25			GAA Glu 230							720
23			AAC Asn						•	768
30			GAA Glu							816
35			GCC Ala							864
40			GCC Ala		TAA				i	894

(2) INFORMATION FOR SEQ ID NO: 4:

5			(SEQU A) L B) T D) T	ENGT YPE:	H: 2 ami	97 a	mino cid			•					
		(ii) MO	LECU	LE T	YPE:	pro	tein	l				•			
10		(xi) SE	QUEN	CE D	BSCR	IPTI	ON:	SEQ	ID N	0:,4	:				
	Met 1		Gln	Leu	Ser 5		Glu	Gln	Phe	Leu 10		Glu	Gln	Lys	Gln 15	
15	Val	Glu	Thr	Ala 20		Ser	Arg	Tyr	Ile 25		Arg	Leu	Glu	Gly 30	Pro	Ala
	Lys	Val	Lys 35		Ala	Met	Ala	Tyr 40		Leu	Glu	Ala	Gly 45	Gly	Lys	Arg
20	Ile	Arg 50	Pro	Leu	Leu	Leu	Leu 55	Ser	Thr	Val	Gln	Ala 60		Gly	Lys	Asp
	Pro 65	Ala	Val	Gly	Leu	Pro 70	Val	Ala	Cys	Ala	11e 75	Glu	Met	Ile	His	Thr 80
25	Tyr	Ser	Leu	Ile	His 85	Asp	Asp	Leu	Pro	Ser 90	Met	Asp	Asn	qeA	As p 95	Leu
	Arg	Arg	Gly	Lys 100	Pro	Thr	Asn	His	Lys 105	Val	Phe	Gly	Glu	Ala 110	Met	Ala
30	Ile	Leu	Ala 115	Gly	As p	Gly	Leu	Leu 120	Thr	Tyr	Ala	Phe	Gln 125	Leu	Ile	Thr
	Glu	Ile 130	Ąsp	Asp	Glu	Arg	Ile 135	Pro	Pro	Ser	Val	Arg 140	Leu	Arg	Leu	Ile
35	Glu 145	Arg	Leu	Ala	Lys	Ala 150	Al'a	Gly	Pro	Glu	Gly 155	Met	Val	Ala	Gly	Gln 160
40	Ala	Ala	Asp	Met	Glu 165	Gly	Glu	Gly	Lys	Thr 170	Leu	Thr	Leu	Ser	Glu 175	Leu
40	Glu	Tyr	Ile	His 180	Arg	His	Lys	Thr	Gly 185	Lys	Met	Leu	Gln	Tyr 190	Ser	Val
45	His	Ala	Gly 195	Ala	Leu	Ile	Gly	Gly 200	Ala	Asp	Ala	Arg	Gln 205	Thr	Arg	Glu
	Leu	Asp 210	Glu	Phe	Ala	Ala	His 215	Leu	Gly	Leu	Ala	Phe 220	Gln	Ile	Arg	Asp
50	Asp 225	Ile	Leu	qaA	Ile	Glu 230	Gly	Ala	Glu	Glu	Lys 235	Ile	Gly	Lys	Pro	Val 240
	Gly	Ser	Asp	Gln	Ser 245	Asn	Asn	Lys	Ala	Thr 250	Tyr	Pro	Ala	Leu	Leu 255	Ser
•																

55

	Leu	Ala	Gly	Ala 260	Lys	Glu	Lys	Leu	Ala 265	Phe	His	Ile	Glu	Ala 270	Ala	Gln
5	Arg	His	Leu 275	Arg	Asn	Ala	Asp	Val 280	Asp	Gly	Ala	Ala	Leu 285	Ala	Tyr	Ile
10	Cys	Glu 290	Leu	Val	Ala	Ala	Arg 295	Asp	His							
15																
20																
<i>25</i>																
30																
35																
40																
45																
50																
<i>55</i>																

	(2)	INE	ORMA	TION	FOF	SE(Q ID	NO:	5:								
5		()	((A) I (B) I (C) S	LENGT TYPE : TRAN	TH: 8 nuc IDEDN	ACTER 394 h cleic NESS:	ase aci	pair d	s							
10		(vi		RIGIN			CE: Bac	illu	s st	earo	ther	moph	ilus	ı			
15		(ix	(B) L	AME/ OCAT THER	ION:	CDS 1 ORMA for	894 TION	: /f nesy	unct ldip	ion= hosp	"Mu hate	tant syn	(3) thas	of e"	DNA	
		(xi) SE	QUEN	CR D	ESCR	IPTI	ON:	SEQ	ID N	0: 5	:					
20	ATG Met	Ala	CAG Gln	CTT Leu	TCA Ser	GTT Val	GAA Glu	CAG Gln	TTT Phe	CTC Leu 10	AAC Asn	GAG Glu	CAA Gln	AAA Lys	Gln	GCG Ala	. 48
				000											15		
25	Val	Glu	Thr	Ala 20	Leu	Ser	Arg	TAT	Ile 25	GAG Glu	Arg	TTA Leu	GAA Glu	GGG Gly 30	CCG Pro	GCG Ala	96
30	AAG Lys	CTG Leu	AAA Lys 35	AAG Lys	GCG Ala	ATG Met	GCG Ala	TAC Tyr 40	TCA Ser	TTG Leu	GAG Glu	GCC Ala	GGC Gly 45	GGC Gly	AAA Lys	CGA A rg	144
	ATC Ile	CGT Arg 50	CCG Pro	TTG Leu	CTG Leu	CTT Leu	CTG Leu 55	TCC Ser	ACC Thr	GTT Val	CGG Arg	GCG Ala 60	CTC Leu	GGA Gly	AAA Lys	GAC Asp	192
35	CCG Pro 65	GCG Ala	GTC Val	GGA Gly	TTG Leu	CCC Pro 70	GTC Val	GCC Ala	TGC Cys	GCG Ala	ATT Ile 75	GAA Glu	ATG Met	ATC Ile	CAT His	ACG Thr 80	24 0
40	TAC Tyr	TCT Ser	TTG Leu	ATC Ile	CAT His 85	GAT Asp	GAT Asp	TTG Leu	CCG Pro	AGC Ser 90	ATG Met	GAC Asp	AAC Asn	GAT Asp	GAT Asp 95	TTG Leu	288
45	CGG Arg	CGC Arg	GGC Gly	AAG Lys 100	CCG Pro	ACG Thr	AAC Asn	CAT His	AAA Lys 105	GTG Val	TTC Phe	GGC Gly	GAG Glu	GCG Ala 110	ATG Met	GCC Ala	336
	ATC Ile	TTG Leu	GCG Ala 115	GGG Gly	GAC Asp	GGG Gly	TTG Leu	TTG Leu 120	ACG Thr	TAC Tyr	GCG Ala	TTT Phe	CAA Gln 125	TTG Leu	ATC Ile	ACC Thr	384
50	GAA Glu	ATC Ile 130	GAC Asp	GAT Asp	GAG Glu	CGC Arg	ATC Ile 135	CCT Pro	CCT Pro	TCC Ser	GTC Val	CGG Arg 140	CTT Leu	CGG Arg	CTC Leu	ATC Ile	432

	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GCC	GCC	GGT	CAG	480
•	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	Gly	Met	Ala	Āla	Gly	Gln	
_	145					150					155				_	160	
5																	
	GCA	GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	528
	Ala	Ala	Asp	Met	Glu	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu	
					165					170					175		
10	C A A	ጥእር	N errer	CATE	000	mam.		3.00	000			~=-					
	Glu	Tur	Tle	Uic	7~~	TAT	AAA	ACC.	GGG	AAA	ATG	CIG	CAA	TAC	AGC	GTG	576
	GIU	TYL	116	180	ALG	Tyr	гу	IIII		ьys	met	rea	GIN		Ser	Val	
				100					185					190			
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	624
15	His	Ala	Gly	Ala	Leu	Ile	Gly	Gly	Ala	Asp	Ala	Arg	Gln	Thr	Ara	Glu	024
			195				•	200			,	3	205		5		
	CTT	GAC	GAA	TTC	GCC	GCC	CAT	CTA	GGC	CTT	GCC	TTT	CAA	ATT	CGC	GAT	672
	Leu	Asp	Glu	Phe	Ala	Ala	His	Leu	Gly	Leu	Ala	Phe	Gln	Ile	Arg	Asp	
20		210					215					220					
																•	
						GAA											720
		Ile	Leu	Asp	Ile	Glu	Gly	Ala	Glu	Glu		Ile	Gly	Lys	Pro	Val	
	225					230					235					240	
25	CCC	NCC.	CNC	CAA	3.00	220											
	Gly	Ser	DAC	CAA	AGC	AAC	AAC	AAA	GCG	ACG	TAT	CCA	GCG	TTG	CTG	TCG	768
	OLY	Jer	лэр	GIII	245	Asn	ASII	пув	MIG	250	ıyı	Pro	Ara	ren		ser	
					243					250					255		
30	CTT	GCC	GGC	GCA	AAG	GAA	AAG	TTG	GCG	TTC	CAT	ATC	GAG	GCG	GCG	CAG	816
	Leu	Ala	Gly	Ala	Lys	Glu	Lys	Leu	Ala	Phe	His	Ile	Glu	Ala	Ala	Gln	010
				260	-		-		265					270			
	CGC	CAT	TTA	CGG	AAC	GCC	GAC	GTT	GAC	GGC	GCC	GCG	CTC	GCC	TAT	ATT	864
35	Arg	His		Arg	Asn	Ala	qaA	Val	Asp	Gly	Ala	Ala	Leu	Ala	Tyr	Ile	
			275					280					285				
	m~~	~~	~~~							_							•
						GCC				AAT							894
			ьeu	vaı	ATS	Ala	_	Asp	His								
40		290					295										

.

(2) INFORMATION FOR SEQ ID NO: 6:

5	(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 297 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein															
		(ii) MO	LECU	LE T	YPE:	pro	tein								
10		(xi) SE	QUEN	CE D	BSCR	IPTI	ON:	SEQ	ID N	o: 6	:				
	Met 1	Ala	Gln	Leu	Ser 5	Val	Glu	Gļn	Phe	Leu 10	Asn	Glu	Gln	Lys	Gln 15	Ala
15	Val	Glu	Thr	Ala 20	Leu	Ser	Arg	Tyr	Ile 25	Glu	Arg	Leu	Glu	Gly 30	Pro	Ala
	Lys	Leu	Lys 35	Lys	Ala	Met	Ala	Tyr 40	Ser	Leu	Glu	Ala	Gly 45	Gly	Lys	Arg
20	Ile	Arg 50	Pro	Leu	Leu	Leu	Leu 55	Ser	Thr	Val	Arg	Ala 60	Leu	Gly	Lys	Asp
	Pro 65	Ala	Val	Gly	Leu	Pro 70	Val	Ala	Cys	Ala	Ile 75	Glu	Met	Ile	His	Thr 80
25	Tyr	Ser	Leu	Ile	His 85	Asp	qeA	Leu	Pro	Ser 90	Met	Asp	Asn	Asp	Asp 95	Leu
	Arg	Arg	Gly	Lys 100	Pro	Thr	Asn	His	Lys 105	Val	Phe	Gly	Glu	Ala 110	Met	Ala
30	Ile	Leu	Ala 115	Gly	Asp	Gly	Leu	Leu 120	Thr	Tyr	Ala	Phe	Gln 125	Leu	Ile	Thr
35	Glu	Ile 130	Asp	Asp	Glu	Arg	Ile 135	Pro	Pro	Ser	Val	Arg 140	Leu	Arg	Leu	Ile
	Glu 145	Arg	Leu	Ala	Lys	Ala 150	Ala	Gly	Pro	Glu	Gly 155	Met	Ala	Ala	Gly	Gln 160
10	Ala	Ala	Asp	Met	Glu 165	Gly	Glu	Gly	Lys	Thr 170	Leu	Thr	Leu	Ser	Glu 175	Leu
	Glu	Tyr	Ile	His 180	Arg	Tyr	Lys	Thr	Gly 185	Lys	Met	Leu	Gln	Tyr 190	Ser	Val
15	His	Ala	Gly 195	Ala	Leu	Ile		Gly 200	Ala	Asp	Ala	Arg	Gln 205	Thr	Arg	Glu
·.	Leu	Asp 210	Glu	Phe	Ala	Ala	His 215	Leu	Gly	Leu	Ala	Phe 220	Gln	Ile	Arg	As p
50 -	Asp 225	Ile	Leu	qeA	Ile	Glu 230	Gly	Ala	Glu	Glu	Lys 235	Ile	Gly	Lys	Pro	Val 240
	Gly	Ser	Asp	Gln	Ser 245	Asn	Asn	Lys	Ala	Thr 250	Tyr	Pro	Ala	Ĺeu	Leu 255	Ser

	Leu	Ala	Gly	Ala 260	Lys	Glu	Lys	Leu	Ala 265	Phe	His	Ile	Glu	Ala 270	Ala	Gln
5	Arg	His	Leu 275	Arg	Asn	Ala	Asp	Val 280	Asp	Gly	Ala		Leu 285	Ala	Tyr	Ile
10	Cys	Glu 290	Leu	Val	Ala	Ala	Arg 295	Asp	His							
15																
20																
25																
30																
35																
40			٥													
45																
50		•														
<i>55</i>																

	ν,		0.0.2				, 10	140.	<i>,</i> .							-	
5		(i	((A) I (B) I (C) S	ICE C ENGT YPE: TRAN	H: 8 nuc DBDN	94 b leic ESS:	ase aci dou	pair d	s							
10		(vi			AL S RGAN			illu	s st	earo	ther	moph	ilus				
15		(ix	(B) L	AME/ OCAT THER	ION: INF	1	894 TION							of : e"	DNA	
20	ATG Met	GCG Ala	CAG	CTT	CE DI TCA Ser 5	GTT	GAA	CAG	TTT	CTC	AAC	GAG	CAA Gln	AAA Lys	CAG Gln 15	GCG Ala	4.8
25	GTG Val	G AA Glu	ACA Thr	GCG Ala 20	CTC	TCC Ser	CGT A rg	TAT Tyr	ATA Ile 25	GAG	CGC Arg	TTA Leu	GAA Glu	GGG Gly 30	CCG Pro	GCG Ala	96
30	AAG Lys	CTG Leu	AAA Lys 35	AAG Lys	GCG Ala	ATG Met	GCG Ala	TAC Tyr 40	TCA Ser	TTG Leu	GAG Glu	GCC Ala	GGC Gly 45	GGC Gly	AAA Lys	CGA Arg	144
	ATC Ile	CGT Arg 50	CCG Pro	TTG Leu	CTG Leu	CTT Leu	CTG Leu 55	TCC Ser	ACC Thr	GTT Val	CGG Arg	GCG Ala 60	CTC Leu	GGC Gly	AAA Lys	GAC Asp	192
35	CCG Pro 65	GCG Ala	GTC Val	GGA Gly	TTG Leu	CCC Pro 70	GTC Val	GCC Ala	TGC Cys	GCG Ala	ATT Ile 75	GAA Glu	ATG Met	ATC Ile	CAT His	ACG Thr 80	240
40															GAT Asp 95		288
45	CGG Arg	CGC Arg	GGC Gly	AAG Lys 100	CCG Pro	ACG Thr	AAC Asn	CAT His	AAA Lys 105	GTG Val	TTC Phe	GGC Gly	GAG Glu	GCG Ala 110	ATG Met	GCC Ala	336
	ATC Ile	TTG Leu	GCG Ala 115	GGG Gly	GAC Asp	GGG Gly	TTG Leu	TTG Leu 120	ACG Thr	TAC Tyr	GCG Ala	TTT Phe	CAA Gln 125	TTG Leu	ATC Ile	ACC Thr	384
50	GAA Glu	ATC Ile 130	GAC Asp	GAT Asp	GAG Glu	CGC Arg	ATC Ile 135	CCT Pro	CCT Pro	TCC Ser	GTC Val	CGG Arg 140	CTT Leu	CGG Arg	CTC Leu	ATC Ile	432

5			AAA Lys							4	80
			GAA Glu 165							5	28
10			CGG Arg							. 5	76
15			TTG Leu							6.	24
20			GCC Ala							6	72
25			ATT Ile							7:	20
			AGC Ser 245							7	68
30			AAG Lys							8:	16
35			AAC Asn							8	64
40			GCC Ala	 	 	TAA				8:	94

(2) INFORMATION FOR SEQ ID NO: 8:

5	(A) LENGTH: 297 amino acids (B) TYPE: amino acid (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein															
10		(xi)	SE	QUEN	CB DI	ESCR:	[PTI	ON: S	SEQ :	ID NO	D: 8	:	•			
M	1et	Ala	Gln	Leu	Ser 5	Val	Glu	Gln	Phe	Leu 10	Asn	Glu	Gln	Lys	Gln 15	Ala
<i>15</i> V	/al	Glu	Thr	Ala 20	Leu	Ser	Arg	Tyr	Ile 25	Glu	Arg	Leu	Glu	Gly 30	Pro	Ala
L	ys	Leu	Lys 35	Lys	Ala	Met	Ala	Tyr 40	Ser	Leu	Glu	Ala	Gly 45	Gly	Lys	Arg
20 I	le	Arg 50	Pro	Leu	Leu	Leu	Leu 55	Ser	Thr	Val	Arg	Ala 60	Leu	Gly	Lys	Asp
	ro 65	Ala	Val	Gly	Leu	Pro 70	Val	Ala	Cys	Ala	Ile 75	Glu	Met	Ile	His	Thr 80
25 H	lis	Ser	Leu	Ile	His 85	Asp	Asp	Leu	Pro	Ser 90	Met	Asp	Asn	Asp	Аз р 95	Leu
А	rg .	Arg	Gly	Lys 100	Pro	Thr	Asn	His	Lys 105	Val	Phe	Gly	Glu	Ala 110	Met	Ala
30 I	le		Ala 115	Gly	Asp	Gly	Leu	Leu 120	Thr	Tyr	Ala		Gln 125	Leu	Ile	Thr
		Ile 130	Asp	Asp	Glu	Arg	Ile 135	Pro	Pro	Ser	Val	Arg 140	Leu	Arg	Leu	Ile
	lu . 45	Arg	Leu	Ala	Lys	Ala 150	Ala	Gly		Glu	Gly 155	Met	Val	Ala	Gly	Gln 160
· A	la i	Ala	Ąsp	Met	Glu 165	Gly	Glu	Gly	Lys	Thr 170	Leu	Thr	Leu	Ser	Glu 175	Leu
· -	lu'	Tyr		His 180	Arg	His	Lys	Thr	Gly 185	Lys	Met	Leu	Gln	Tyr 190	Ser	Val
н <i>5</i>	is A	_	_			Ile	_	-		_		_			Arg	Glu
L		Asp 210	Glu	Phe	Ala	Ala	His 215	Leu	Gly	Leu	Ala	Phe 220	Gln	Ile	Arg	Asp
	sp 25	Ile:	Leu .	Asp		G1u 230	Gly	Ala	Glu	Glu	Lys 235	Ile	Gly	Lys	Arg	Val 240
G.	ly s	Ser .	Asp		Ser 245	A sn	Asn	Lys	Ala	Thr 250	Tyr	Pro	Ala	Leu	Leu 255	Ser

	Leu	Ala	GIA	A1a 260	Lys	Glu	Lys	Leu	Thr 265	Phe	His	Ile	Glu	Ala 270	Ala	Gln
5	Arg	His	Leu 275	Arg	Asn	Ala	Asp	Val 280	Asp	Gly	Ala	Ala	Leu 285	Ala	Tyr	Ile
10	Cys	Glu 290	Leu	Val	Ala	Ala	Arg 295	Asp	His							
15								·								
20																
25																
30															in .	
35																
40																
45																
50																
55																

(2) INFORMATION FOR SEQ ID NO: 9:

55

5		(i)	((A) I (B) T (C) S	ENGT TYPE : TRAN	TH: 8 nuc IDBDN	ACTER 194 b :leic !BSS: lin	ase aci dou	pair d	s						·		
10		(vi			IAL S		E: Bac	illu	s st	earo	ther	moph	ilus	I				
15		(ix	(B) L	AME/ OCAT THER	ION:	CDS 1 ORMA yldi	894 TION	: /f phat	unct e sy	ion= ntha	"DN	A co	ding	for	nativ	⁄e	
		(xi) SE	QUEN	CE D	ESCR	IPTI(ON:	SEQ	ID N	0: 9	:						
20	ATO Met	GCG Ala	C A G	CTT	TCA Ser 5	Val	GAA Glu	CAG Gln	TTT Phe	CTC Leu 10	AAC Asn	GAG Glu	CAA Gln	AAA Lys	CAG Gln 15	Ala		48
25	GTG Val	GAA Glu	ACA Thr	GCG Ala 20	CTC Leu	TCC Ser	CGT Arg	TAT Tyr	ATA Ile 25	GAG Glu	CGC Arg	TTA Leu	GAA Glu	GGG Gly 30	CCG Pro	GCG Ala		96
30	AAG Lys	CTG Lys	AAA Lys 35	AAG Lys	GCG Ala	ATG Met	GCG Ala	TAC Tyr 40	TCA Ser	TTG Leu	GAG Glu	GCC Ala	GGC Gly 45	GGC Gly	AAA Lys	CGA Arg		144
	ATC Ile	CGT Arg 50	CCG Pro	TTG Leu	CTG Leu	CTT Leu	CTG Leu 55	TCC Ser	ACC Thr	GTT Val	CAG Gln	GCG Ala 60	CTC Leu	GGC Gly	AAA Lys	GAC Asp		192
35	CCG Pro 65	GCG Ala	GTC Val	GGA Gly	TTG Leu	CCC Pro 70	GTC Val	GCC Ala	TGC Cys	GCG Ala	ATT Ile 75	GAA Glu	ATG Met	ATC Ile	CAT His	ACG Thr 80		240
40	TAC Tyr	TCT Ser	TTG Leu	ATC Ile	CAT His 85	GAT Asp	GAT Asp	TTG Leu	CCG Pro	AGC Ser 90	ATG Met	GAC Asp	AAC Asn	GAT Asp	GAT Asp 95	TTG Leu		288
45	CGG A rg	CGC Arg	GGC Gly	AAG Lys 100	CCG Pro	ACG Thr	AAC Asn	CAT His	AAA Lys 105	GTG Val	TTC Phe	GGC Gly	GAG Glu	GCG Ala 110	ATG Met	GCC Ala	•	336
	ATC Ile	TTG Leu	GCG Ala 115	GGG Gly	GAC Asp	GGG Gly	Leu	TTG Leu 120	ACG Thr	TAC Tyr	GCG Ala	TTT Phe	CAA Gln 125	TTG Leu	ATC Ile	ACC Thr		384
50	GAA Glu	ATC Ile 130	GAC Asp	GAT Asp	GAG Glu	CGC Arg	ATC Ile 135	CCT Pro	CCT Pro	TCC Ser	GTC Val	CGG Arg 140	CTT Leu	CGG Arg	CTC Leu	ATC Ile		432

	GAA	CGG	CTG	GCG	AAA	GCG	GCC	GGT	CCG	GAA	GGG	ATG	GTC	GCC	GGT.	CAG	48	0
	Glu	Arg	Leu	Ala	Lys	Ala	Ala	Gly	Pro	Glu	Gly	Met	Val	Ala	Gly	Gln		
	145					150					155					160		
5																		
	GCA	GCC	GAT	ATG	GAA	GGA	GAG	GGG	AAA	ACG	CTG	ACG	CTT	TCG	GAG	CTC	52	8
	Ala	Ala	Asp	Met	Glu	Gly	Glu	Gly	Lys	Thr	Leu	Thr	Leu	Ser	Glu	Leu		
					165					170					175			
10	GAA	TAC	ATT	CAT	CGG	CAT	AAA	ACC	GGG	AAA	ATG	CTG	CAA	TAC	AGC	GTG	57	6
	Glu	Tyr	Ile	His	Arg	His	Lys	Thr	Gly	Lys	Met	Leu	Gln	Tyr	Ser	Val		
				180					185					190				
	CAC	GCC	GGC	GCC	TTG	ATC	GGC	GGC	GCT	GAT	GCC	CGG	CAA	ACG	CGG	GAG	62	4
15					Leu													_
			195			_		200		•			205					
					GCC												67	2
•	Leu	_	Glu	Phe	Ala	Ala		Leu	Gly	Leu	Ala		Gln	Ile	Arg	qaA		
20		210					215					220						
	GAT	АТТ	CTC	GAT	ATT	GAA	GGG	GCA	GAA	CAA	מממ	ΔΤΟ	GGC	AAG	CCG	CTC	72	Λ
					Ile													•
	225					230	4- 1				235		4 23	2,0		240		
25																		
	GGC	AGC	GAC	CAA	AGC	AAC	AAC	AAA	GCG	ACG	TAT	CCA	GCG	TTG	CTG	TCG	76	8
	Gly	Ser	Asp	Gln	Ser	Asn	Asn	Lys	Ala	Thr	Tyr	Pro	Ala	Leu	Leu	Ser		
					245					250					255			
00	CTT	GCC	GGĊ	GCG	AAG	AAD	באמ	ጥጥር፤	GCG	ጥጥር	ሮልጥ	ልጥሮ	GAG	GCG	ace	CAG	81	_
30					Lys												01	5
			7	260	-,-		_,,		265		*****		010	270	77.4	G 211		
	CGC	CAT	TTA	CGG	AAC	GCC	GAC	GTT	GAC	GGC	GCC	GCG	CTC	GCC	TAT	ATT	86	4
35	Arg	His	Leu	Arg	Asn	Ala	Asp	Val	Asp	Gly	Ala	Ala	Leu	Ala	Tyr	Ile		
			275			•		280					285					
	TCC	GNA	CTC	מייר	GCC	ccc	ccc	CNC	C እ m	TD 3.3							0.0	
					Ala					TAA							894	#
	-73	290	264	· a I		WT.	295	usb	1112									
40																		

(2) INFORMATION FOR SEQ ID NO: 10:

5			(A) L B) T	ENCE ENGT YPE: OPOL	H: 2 ami	97 a no a	mino cid								
10					LE T		-					_				
10	••••				CE D				_							
	Met 1		Gin	Leu	Ser 5		GIu	Gln	Phe	Leu 10	Asn	Glu	Gln	Lys	Gln 15	Ala
15	Val	Glu	Thr	Ala 20		Ser	Arg	Tyr	Ile 25	Glu	Arg	Leu	Glu	Gly 30		Ala
	Lys	Lys	Lys 35	Lys	Ala	Met	Ala	Tyr 40	Ser	Leu	Glu	Ala	Gly 45	Gly	Lys	Arg
20	Ile	Arg 50	Pro	Leu	Leu	Leu	Leu 55	Ser	Thr	Val	Gln	Ala 60	Leu	Gly	Lys	Asp
	Pro 65	Ala	Val	Gly	Leu	Pro 70	Val	Ala	Cys	Ala	Ile 75	Glu	Met	Ile	His	Thr 80
25	Tyr	Ser	Leu	Ile	His 85	Asp	Asp	Leu	Pro	Ser 90	Met	Asp	Asn	Asp	Asp 95	Leu
	Arg	Arg	Gly	Lys 100	Pro	Thr	Asn	His	Lys 105	Val	Phe	Gly	Glu	Ala 110	Met	Ala
30	Ile	Leu	Ala 115	Gly	Asp	Gly	Leu	Leu 120	Thr	Tyr	Ala	Phe	Gln 125	Leu	Ile	Thr
	Glu	Ile 130	Asp	Asp	Glu	Arg	Ile 135	Pro	Pro	Ser	Val	A rg 140	Leu	Arg	Leu	Ile
<i>35</i>	Glu 145	Arg	Leu	Ala	Lys	Ala 150	Ala	Gly	Pro	Glu	Gly 155	Met	Val	Ala	Gly	Gln 160
40	Ala	Ala	Ąsp	Met	Glu 165	Gly	Glu	Gly	Lys	Thr 170	Leu	Thr	Leu	Ser	Glu 175	Leu
	Glu	Tyr	Ile	His 180	Arg	His	Lys	Thr	Gly 185	Lys	Met	Leu	Gln	Tyr 190	Ser	Val
45	His	Ala	Gl y 195	Ala	Lėu	Ile	Gly	Gly 200	Ala	Asp	Ala	Arg	Gln 205	Thr	Arg	Glu
	Leu	Asp 210	Glu	Phe	Ala	Ala	His 215	Leu	Gly	Leu	Ala	Phe 220	Gln	Ile	Arg	Asp
50	Asp 225	Ile	Leu	Asp	Ile	Glu 230	Gly	Ala	Glu	Glu	Lys 235	Ile	Gly	Lys	Pro	Val 240
	Gly	Ser	Asp	Gln	Ser 245	Asn	Asn	Lys	Ala	Thr 250	Tyr	Pro	Ala	Leu	Leu 255	Ser
55																

Leu Ala Gly Ala Lys Glu Lys Leu Ala Phe His Ile Glu Ala Ala Gln 260

Arg His Leu Arg Asn Ala Asp Val Asp Gly Ala Ala Leu Ala Tyr Ile 270

Cys Glu Leu Val Ala Ala Arg Asp His 290

15 Claims

5

10

20

30

40

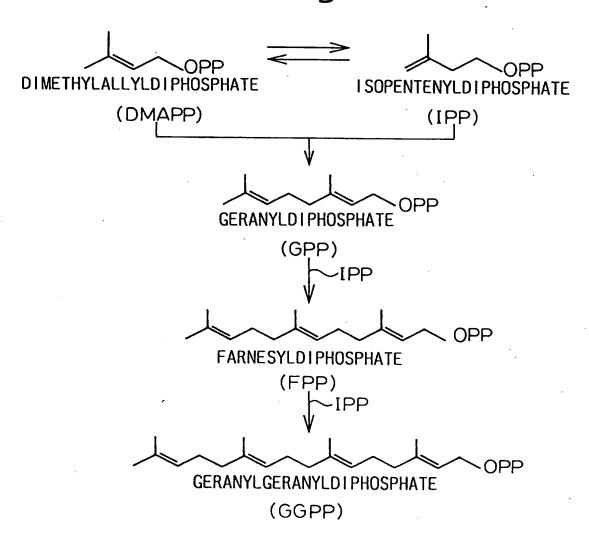
45

50

- 1. A process for production of a gene coding for a mutated farnesyldiphosphate synthase capable of synthesizing geranyldiphosphate synthase comprising the steps of:
 - (1) subjecting a gene coding for a farnesyldiphosphate synthase to mutagenesis;
 - (2) expressing the genes subjected to the mutagenesis; and
 - (3) selecting a gene coding for a mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate.
- A process according to claim 1, wherein the gene coding for mutated farnesyldiphosphate synthase capable of synthesizing geranyldiphosphate is derived from <u>Bacillus stearothermophilus</u>.
 - 3. A gene coding for a mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate synthase, obtainable according to a process of claim 1.
 - 4. A process for production of a mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate, comprising the step of expressing a gene of claim 3.
- 5. A mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate, obtainable according to a process of claim 4.
 - 6. A process for production of geranylgeranyldiphosphate or geranylgeranyol, comprising the step of acting a mutated farnesyldiphosphate synthase capable of synthesizing geranylgeranyldiphosphate on a substrate selected from the group consisting of isopentenyldiphosphate, dimethylallyldiphosphate, geranyldiphosphate and farnesyldiphosphate.
 - 7. A geranylgeranyldiphosphate synthase having an amino acid sequence modified from a native amino acid sequence of a farnesyldiphosphate synthase wherein the modification comprises deletion of one to a few amino acid residues, addition of one to a few amino acid residues or replacement of one to a few amino acid residues with other amino acid residues, or a combination of said modification.
 - 8. A geranylgeranyldiphosphate synthase according to claim 7, wherein the modification is present on at least one of the positions 34, 59, 81, 157, 182, 239, 265 and 275 of farnesyldiphosphate synthase of <u>Bacillus stearothermophilus</u> origin, or one to a few corresponding positions in an amino acid sequence of a farnesyldiphosphate synthase of other origin.
 - 9. A geranylgeranyldiphosphate synthase having an amino acid sequence shown in SEQ ID NO: 1.
 - 10. A geranylgeranyldiphosphate synthase having an amino acid sequence shown in SEQ ID NO: 2.
 - 11. A geranylgeranyldiphosphate synthase having an amino acid sequence shown in SEQ ID NO: 3.
 - 12. A geranylgeranyldiphosphate synthase having an amino acid sequence shown in SEQ ID NO: 4.

13. A gene coding for a geranylgeranyldiphosphate synthase according to claim 7. 14. A gene coding for a geranylgeranyldiphosphate synthase according to claim 8. 5 15. A gene coding for a geranylgeranyldiphosphate synthase according to claim 9. 16. A gene coding for a geranylgeranyldiphosphate synthase according to claim 10. 17. A gene coding for a geranylgeranyldiphosphate synthase according to claim 11. 10 18. A gene coding for a geranylgeranyldiphosphate synthase according to claim 12. 19. An expression vector comprising a gene according to claim 13. 20. An expression vector comprising a gene according to claim 14. 21. An expression vector comprising a gene according to claim 15. 22. An expression vector comprising a gene according to claim 16. 23. An expression vector comprising a gene according to claim 17. 24. An expression vector comprising a gene according to claim 18. 25. A recombinant host transformed with an expression vector according to claim 19. 26. A recombinant host transformed with an expression vector according to claim 20. 27. A recombinant host transformed with an expression vector according to claim 21. 30 28. A recombinant host transformed with an expression vector according to claim 22. 29. A recombinant host transformed with an expression vector according to claim 23. 30. A recombinant host transformed with an expression vector according to claim 24. 40 45 50

Fig.1



CAI EMIHT YSLIHDDLPSMDNDDLRRGKPTN HKVFGEAMAII C I H AYSLIHDDLPAMDDDDLRRGLPTC HVKFGEANAII C I ELLQ AYFLVADD MMDKSITRRGQP C WYKVPEVGEI C VELLQ AFFLVADD IMDSSLTRRGQ TC WYGKPGVGLD C VELLQ AFFLVLDD IMDSSHTRRGQI C WYGKPGIGLD

YAILSNKTVEQLGQEEYEKVAILGW

VRALGKDPAVGLPVA GHMFGVSTNTLDAPAAAVE FRELVEPRKQDADSLQRAWTVGW FQELVEPRKQDAESLQRALTVGW

F. q. 2

⋖

LLLST	LWAT	SWOT	TWWA	TWGT	3	
A YSLEAGGKRIRPL LLLST	Q YGALLGGKRLRPF LVYAT	LNY NTPGGKLNRGL SVVOT	EY NAIGGKYNRGL	EY NTVGGKYNRGL		
MAQL SVEQFLNEQKQAVETALSRY! ERLEGPAKLKKAM		MASEKE! RRERFLNVFPKLVEELNASLLAYGMPKEACDWYAHS	MNGDQNSDVYAQEKQDFVQHFSQIVRVLTEDEMGHPEIGDAIARLKEV LEY NAIGGKYNRGL TVVVA	MNGDQKLDVHNQEKQNFIQHFSQIVKVLTEDELGHPEKGDAITRIKEV LEY NTVGGKYNRGL TVVQT		80
((2)	(3)	(4)	(2)		

EGEGKTLTLSE	DAEGKHVPLDA	GQLMDL I TAPEDKVOLS	CQTLDL LTAPQGNVDLV	CQTLDL ITAPQCQVDLG
GQAADM	GOALDL	GOLMDL	COTLOL	COTLOL
A GDG LLTYA FQLITEIDDERIPPSVRLRLIERLAKAAGPEGMVA GQAADM EGEGKTLTLSE	A GDALQTL A FSILSDADMPEVSDRDRISMISELASASGIAGMCG GQALDL DAEGKHVPLDA	AINDAF ML EA AI YKLLKSHFRNEKYYIDI TELFHEVTFQTEL	AINDAN LL EA CIYRLLKLYCREQPYYLNLIELFLQSSYQTEI	AINDAL LL EA AIYRLLKFYCREQPYYLNLLELFLQSSYQTEI
TYA	A	EA	EA.	- EA
=	9	₹	=	=
A GDG	A GDAL	AINDAF	AINDAN	AINDAL
(1)	(5)	(3)	(4)	(2)

F. q. 3

ш

QTRELDEFAAHL	RRALPVLDKYAESI	DLKQARDVL I PL	RFTEKRYKSIVKY KTAFYSFYLPIAA AMYMAGI D G EKEHANAKKILLEM	RYTEKRYKSIVKY KTAFYSFYLPIAA AMYMAGI D G EKEHANALKILLEM
ADAR	OKG	TOEK	50	<u>ე</u>
ALIG G	SALS AG	AMYVAGI	AMYMAGI	AMYMAGI
KTGKMLQYSVHAG ALIG G ADAR QTRELDEFAAHL	KTGA LIRAAVRLGALS AG OKG RRALPVLOKYAESI	KFSLKKHSFIVTF KTAYYSFYLPVAL AMYVAGITDEK DLKQARDVLIPL	KTAFYSFYLPIAA	KTAFYSFYLPIAA
LEYIHRH	LERIHRH	KFSLKKHSF I VTF	RFTEKRYKSIVKY	RYTEKRYKS! VKY
$\widehat{}$	\sim	3	4	<u></u>

. Ц_

(1) GLAFQIRDDILDIEGAEEKI GKPVGSD QSNNKAT YPALLSLAGAKEKLAFHIEAAQRHLRNADVDGAA (2) GLAFQVQDDILDVVGDTA TLGKRQGAD QQLGK S TYPALLGLEQARKKARDLIDDARQSLKQLAEQSLDTS (3) GEYFQIQDDYLDCFGTPEQI GKI GTDIQDN KCS WVINKALELASAEQRKTLDENYGKKDSVAEAKCKKIF (4) GEFFQIQDDYLDLFGDPSVT GKI GTDIQDN KCS WLVVQCLQRATPEQYQILKENYGQKAEKVARVKALYE (5) GEFFQIQDDYLDLFGDPSVT GKV GTDIQDN KCS WLVVQCLLRATPQQRQILEENYGQKDPEKVARVKALY

ပ

of Higoligalite and a VI)	(1) D. SIEARUINERMUPHILAS	(2) E. COL I	(3) YEASI	(4) HUMAN	(3) KAI
	L AYICELVAARDH	LEA LADYI IORNK	LTAFLN KVYKRSK	LG LARKIYKRK	LE LANKIYKRKK
		V	NOLK!EQLYHEYEES!AKDLKAK!SQVDESRGFKADV LTAFLN KVYKRSK	ELDLPAVFLQYEEDSYSHIMALIEQYAAPLPPAVF LG LARKIYKRRK	EELOLRSVFFKYEEDSYNRLKSLIEQCSAPLPPSIF LE 1
,	`	\sim	3	3	$\overline{\Omega}$

Fig.4

		2 34
W.T	1:	MAQLSVEQFLNEQKQAVETALSRYIERLEGPAKLKKAMAYSLEAGGKRIR
No. 1	1:	
No. 2	1:	V
No. 3	1:	
No.4	1:	
		59 81
W.T	51:	PLLLLSTVRALGKDPAVGLPVACAIEMIHTYSLIHDDLPSMDNDDLRRGK
No.1	51:	H
No. 2	51:	Q
No.3	51:	
No.4	51:	. Н
		141
wr	101 .	141 PTNIKVFGEAMAILAGDGLLTYAFQLITEIDDERIPPSVRLRLIERLAKA
W.T No.1	101:	PINNKYFUEAMAILAGUGLLIIAFULIIEIUUEKIPPSVKLKLIEKLAKA
No. 2	101:	
No. 3	101 :	
No. 4	101:	
110. 4	101 .	
		157 182
W.T	151 :	AGPEGMVAGQAADMEGEGKTLTLSELEYIHRIIKTGKMLQYSVHAGALIGG
No. 1	151 :	
No. 2	151:	
No.3	151 :	A Y
No. 4	151:	
		239
W.T	201:	ADARQTRELDEFAAHLGLAFQIRDDILDIEGAEEKIGKPVGSDQSNNKAT
No.1	201:	
No. 2	201:	
No.3	201:	
No.4	201 :	R
		265 275
W.T	251 :	
No. 1	251:	
No. 2	251 :	
No. 3	251 :	
No.4	251 :	T

Fig.5

